UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

NEW CONCEPT OF IMMUNOTHERAPY FOR GENE MANIPULATION

ANA SOFIA DA SILVA NARCISO

MESTRADO EM CIÊNCIAS BIOFARMACÊUTICAS

2016

DISSERTAÇÃO ORIENTADA PELO PROF. DOUTOR JOÃO GONÇALVES

UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

NEW CONCEPT OF IMMUNOTHERAPY FOR GENE MANIPULATION

ANA SOFIA DA SILVA NARCISO

MESTRADO EM CIÊNCIAS BIOFARMACÊUTICAS

2016

DISSERTAÇÃO ORIENTADA PELO PROF. DOUTOR JOÃO GONÇALVES

“WHAT’S STOPPING YOU?

(NOTHING.)”

UNKNOWN

AOS MEUS PAIS

ABSTRACT

The knowledge of gene regulation led to the emergence of gene therapy as a versatile tool for the prevention and treatment of a variety of human diseases. New strategies for gene manipulation have been developed to date. Due to their innate ability to cross cell membranes and bind DNA in a specific and efficient manner, zinc-finger proteins possess a great potential for gene targeting and regulation. Although, zinc-fingers present the critical limitation of their low specificity to the target cells. Antibody fragments have demonstrated to have high specificity which make them a powerful tool to overcome this limitation presented by zinc-fingers.

Within this context, this thesis aims to study the improvement of a previously studied therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers. To validate our strategy and as a proof-of-concept, we choose HIV-1 as the disease model since Acquired Immune Deficiency Syndrome (AIDS) is a disease which represent a major global public health issue. Therefore, we engineered four alternative bispecific proteins of an artificial zinc-finger (KRAB-HLTR3) designed to repress the transcription from the HIV-1 LTR promoter coupled to a CXCR4-specific nanobody (VHH).

These proteins were successfully expressed and purified with high yields of soluble protein. Afterwards, we evaluated the proteins specificity and affinity to their targets and we demonstrated that these alternative bispecific proteins bind specifically to HLTR3 binding site in a concentration-dependent manner, similarly to previously studied proteins. We also demonstrated that these bispecific proteins bind specifically to CXCR4 receptor at the surface, similarly to previously studied proteins. Finally, we evaluated the bispecific proteins ability to repress transcription from the HIV-1 LTR promoter. We assessed repression of transcription of a destabilized green fluorescent protein (GFP) reporter and we demonstrated that these constructions repress transcription of GFP gene driven by the HIV-1 LTR promoter in a concentration-dependent manner, like to previously studied proteins. Although results suggest that zinc-finger on the C-terminal promote the binding to the HIV-1 LTR promoter and consequently the protein ability to repress viral transcription.

With this in mind, more *in vitro* studies have to be performed to evaluate the bispecific proteins ability to repress transcription of the HIV-1 genome. In fact, since activation of the HIV-1 LTR promoter lead to expression of the HIV-1 genome, these

results suggest that in the presence of the HIV-1 genome, these proteins inhibit the HIV-1 LTR promoter and consequently repress transcription of the HIV-1 genome. Additionally, to evaluate the repression of viral replication by these bispecific proteins, infection assays must be performed in Jurkat cell line or primary $CD4^+$ T-lymphocytes. The infection can be performed with HIV-1 laboratory-adapted strains or HIV-1 primary isolates.

In conclusion, results presented in this dissertation demonstrated that these therapeutic proteins improve the antibody delivery strategy to gene manipulation previously developed in our laboratory. In fact, these proteins are a promising tool to be applied in the clinical and might complement other gene-based strategies. Furthermore, these recombinant proteins can be designed and engineered to use in other therapeutic applications.

Key-Words: Gene therapy; Zinc-fingers proteins; Antibody engineering; Antibody Delivery; Human immunodeficiency virus type 1 (HIV-1).

RESUMO

Nas últimas duas décadas, a terapia génica tem emergido como uma alternativa promissora para o tratamento de uma grande variedade de doenças. Esta terapia consiste na transferência de transgenes para células alvo, promovendo a correção de anormalidades no fenótipo ou genótipo dos pacientes.

O crescente conhecimento da regulação génica, assim como da estrutura e função do genoma humano levaram ao aparecimento de novas estratégias para manipulação da expressão génica. O sucesso destas novas estratégias terapêuticas está intimamente relacionado com a escolha de sistemas de entrega de genes eficientes, específicos e não tóxicos. Até ao momento várias estratégias para manipulação génica foram desenvolvidas e testadas. As mais comuns incluem vetores virais e algumas formas de ADN não viral.

O uso de vetores virais como veículos de transferência e expressão génica representa um poderoso instrumento, dada a capacidade que os vírus apresentam de penetrar dentro do núcleo das células hospedeiras e explorar a sua maquinaria celular. Entre os mais usados estão os vetores retrovirais, adenovirais e adeno-associados que apesentam alta eficiência de transfecção *in vivo*. No entanto, a elevada imunogenicidade, o limitante tamanho do transgene e ainda a produção de toxinas são alguns problemas destas estratégias.

Relativamente às estratégias não virais, elas incluem inoculação de ADN puro ou encapsulado através de técnicas como microinjeção e electroporação. Os sistemas não virais representam uma importante alternativa aos virais, dada a sua menor imunogenicidade e não apresentarem limitações no tamanho do transgene. Apesar das suas vantagens, estes vetores apresentam uma transfecção menos eficiente comparativamente aos virais.

Com o intuito de ultrapassar algumas limitações, quer dos vetores virais como dos não virais, as proteínas dedos de zinco (do inglês, *zinc-finger proteins (ZFs))* surgiram como ferramentas versáteis para a terapia génica. Os ZFs apresentam uma capacidade inata para atravessar as membranas celulares e de se ligarem eficientemente e especificamente ao ADN. Desta forma, os ZFs podem ser desenhados para reconhecer uma vasta gama de sequência de ADN de modo a ativar, reprimir,

cortar ou colar genes. No entanto, apresentam ainda algumas limitações que devem ser ultrapassadas no futuro, em particular a sua baixa especificidade para as células alvo.

Novas estratégias terapêuticas têm sido desenvolvidas de modo a ultrapassar esta limitação dos ZFs. De facto, dada a sua alta especificidade, os anticorpos monoclonais bem como pequenos derivados de anticorpos recombinantes têm demonstrado um enorme potencial para combater este problema dos ZFs.

Neste contexto, este projeto científico tem como objetivo melhorar uma estratégia terapêutica de manipulação génica através da entrega de ZFs por anticorpos anteriormente desenvolvida no nosso laboratório. Como prova do conceito, escolhemos como modelo de doença o síndroma de imunodeficiência adquirida (SIDA), uma doença infeciosa de incidência mundial causada pelo vírus da imunodeficiência humana (VIH-1). A infeção pelo VIH-1 é caracterizada por uma supressão do sistema imunitário, levando ao aparecimento de doenças oportunistas. Este vírus tem a capacidade de infetar células $CD4^+$ como é o caso dos linfócitos T e dos macrófagos. No entanto, não só o receptor celular CD4 permite a ligação e consequente entrada do vírus nas células alvo, o receptor de quimiocinas CXCR4 como também o CCR5 determinam também o tropismo celular do vírus. Apesar dos progressos realizados no tratamento da SIDA, especialmente através do uso de fármacos antirretrovirais (HAART), estes não são capazes de erradicar por completo o vírus do organismo. Deste modo, o desenvolvimento de novas estratégias contra o VIH-1, tal como a terapia génica, mostra especial interesse.

Com este propósito, foram inicialmente construídas quatro proteínas biespecíficas variantes das já desenvolvidas no nosso laboratório, compostas por um anticorpo recombinante (VHH, também designado por *nanobody*) desenhado contra o receptor CXCR4 e um ZF com o domínio repressor *KRAB* desenhado para reprimir a transcrição do genoma do VIH-1 através da sua ligação ao promotor LTR do VIH-1 (KRAB-HLTR3). De forma a avaliar qual a conformação que proporciona uma maior estabilidade e solubilidade à proteína, uma das construções foi desenhada com o KRAB-HLTR3 a N-terminal, similarmente a uma das proteínas já desenvolvidas, enquanto nas outras três foi adicionado a C-terminal. Além disso, para facilitar a libertação do ZF nas células alvo, em duas construções foi introduzida a sequência de clivagem da catepsina B, uma cisteína proteinase lisossomal. Com o mesmo propósito, numa das construções foi introduzida a sequência de clivagem da MMP-9. Por outro lado, foi feita a construção de uma proteína que consiste apenas no KRAB-HLTR3 que foi gentilmente cedida pela C. Cunha-Santos (Laboratório João Gonçalves).

Após a construção, todas as proteínas recombinantes foram clonadas no mesmo vector de expressão bacteriano, expressas em *E. coli* e posteriormente purificadas. Com exceção da construção que possui o sítio de clivagem da MMP-9, todas as outras foram purificadas com sucesso, sendo utilizadas nos ensaios seguintes.

Seguidamente, para avaliar a capacidade de ligação de cada proteína biespecífica ao promotor LTR do VIH-1, foram realizados ensaios preliminares de ligação por ELISA, usando uma sequência de oligonucleótidos reconhecida pelo KRAB-HLTR3 como antigénio (sitio de ligação do HLTR3, do inglês *HLTR3 binding site*). Verificámos que estas proteínas recombinantes ligam especificamente ao sítio de ligação do HLTR3 e que esta ligação é dependente da concentração, o que se assemelha com as proteínas anteriormente estudadas.

Sendo que as proteínas são biespecíficas foi necessário avaliar a funcionalidade dos dois domínios funcionais. Com este propósito, um ensaio de citometria de fluxo na linha celular Jurkat E6-1 T foi realizado e verificámos que estas proteínas recombinantes ligam especificamente à superfície do receptor CXCR4, similarmente ao observado com as proteínas já estudadas.

De forma a validar a capacidade das proteínas biespecíficas de reprimir a transcrição viral através do promotor LTR do VIH-1, foi avaliada a repressão da transcrição de um gene repórter, nomeadamente a GFP (do inglês *green fluorescent protein*) conduzido pelo promotor LTR do VIH-1. Ensaios de citometria de fluxo foram realizados na linha celular HeLa-Tat-III/LTR/d1EGFP e verificámos que estas proteínas biespecíficas reprimem a transcrição do gene da GFP e que esta repressão é dependente da concentração, o que se assemelha com as proteínas estudadas anteriormente. Verificámos no entanto que as construções que possuem o KRAB-HLTR3 a C-terminal apresentam uma maior capacidade de repressão da transcrição. Uma vez que a ativação do promotor LTR do VIH-1 leva à expressão do genoma do VIH-1, estes resultados sugerem que na presença do genoma do VIH-1 é possível inibir o promotor LTR e consequentemente reprimir a transcrição do genoma do VIH-1.

Simultaneamente, e dado ao facto dos fragmentos de anticorpos exibirem algumas limitações farmacocinéticas, outra estratégia terapêutica foi desenvolvida. Desta forma, construímos três proteínas biespecíficas compostas por um anticorpo monoclonal desenhado contra o receptor HER2 (Trastuzumab, Herceptin®) acoplado a um ZF com o domínio repressor *KRAB* desenhado para se ligar ao promotor do protooncogene *erbB-2/HER-2.* Para facilitar a libertação do ZF nas células alvo, numa das construções foi introduzida a sequência de clivagem da catepsina B e noutra

construção a sequência de clivagem da MMP-9. Os resultados dos ensaios preliminares de transfecção mostraram que à exceção da construção que possui a sequência da MMP-9, as restantes duas foram construídas e expressas na linha celular HEK293T com sucesso.

Em conclusão, os resultados apresentados neste projeto científico demonstraram que estas proteínas biespecíficas melhoram a estratégia terapêutica de manipulação génica anteriormente desenvolvida no nosso laboratório. De facto, estas proteínas biespecíficas apresentam um enorme potencial para serem aplicadas na clinica.

Como perspetivas futuras, mais ensaios de citometria de fluxo na linha celular Jurkat E6-1 T são necessários de forma a avaliar a internalização das proteínas biespecíficas via CXCR4. Além disso, para validar que a ligação e a internalização das proteínas ocorre efetivamente via CXCR4, os mesmos ensaios terão de ser realizados na linha celular Jurkat CXCR4 negativas. Por outro lado, para avaliar a influência da catepsina B na libertação do ZF, ensaios funcionais deverão ser realizados. Relativamente à capacidade das proteínas biespecíficas de reprimirem a transcrição do genoma do VIH-1 através da inibição do promotor LTR, terão de ser realizados ensaios *in vitro* numa linha celular que integre o genoma do VIH-1. Adicionalmente para avaliar a repressão da replicação viral por estas proteínas biespecíficas, ensaios de infeção deverão ser feitos na linha celular Jurkat ou em linfócitos primários.

Palavras-chave: Terapia génica; Proteínas dedos de zinco; Engenharia de anticorpos; Entrega de anticorpos; Vírus da imunodeficiência humana tipo 1 (VIH-1).

AKNOWLEDGMENTS

Em primeiro lugar, gostaria de agradecer ao Professor Doutor João Gonçalves, orientador desta tese. Agradeço-lhe pela oportunidade que me deu e pelo seu voto de confiança. Agradeço-lhe por todos os conhecimentos transmitidos e por toda a paciência que me dispensou. Muito obrigada por me ter acompanhado neste caminho nem sempre fácil, pela amizade, ajuda e apoio ao longo deste ano em que fiz parte da sua equipa de investigação.

Quero agradecer aos meus colegas de laboratório, em especial à Catarina Santos, à Ana Catarina Santos, ao Pedro Perdigão, à Paula, e também aos antigos colegas, ao Renato, à Rita e à Sofia. Obrigada pela disponibilidade dada e paciência que tiveram para partilhar comigo tantos conhecimentos, e pelo companheirismo que foi essencial para a minha integração no grupo.

Quero fazer um agradecimento especial à Joana Oliveira por partilhar comigo esta viagem. Sem ti não teria sido possível. Muito obrigada pelos dias e noites de trabalho, de apoio, de partilha, de amizade e companheirismo!

Agradeço ainda à Inês Lourenço por me ensinar a definição de Amizade. Obrigada por me apoiares e me acompanhares sempre que eu precisava de uma pausa para carregar baterias. Obrigada pelas nossas loucuras e grandes momentos!

Por último, mas em primeiro no meu coração, quero fazer um agradecimento muito especial aos meus pais por sempre terem acreditado em mim e me terem dado a oportunidade de ser o que sou hoje. Muito obrigada pelo vosso apoio incondicional, pela força e por estarem sempre ao meu lado! Obrigada Mãe pelas gargalhadas e lágrimas partilhadas, pelos desabafos e birras que aturaste, pelo amor e carinho que sempre me dedicaste! Obrigada Pai por todos os sonhos que incentivaste, por todo o amor, educação e princípios que me proporcionaste! Muito obrigada pais por todo o orgulho que demonstram ter em mim!

ABBREVIATIONS

GENERAL

Abs	Absorbance
ADCC	Antibody Dependent Cell Mediated Cytotoxicity
AIDS	Acquired Immunodeficiency Syndrome
C_H	Constant Heavy Chain
C_L	Constant Light Chain
CXCR4	CXC Chemokine Receptor 4
DNA	Deoxyribonucleic Acid
E. coli	*Escherichia coli*
Fab	Fragment Antigen Binding
Fc	Fragment Crystallizable
FcRn	Neonatal Fc Receptor
Fv	Fraction Variable
h	hour
HA	Hemagglutinin tag
HAART	Highly Active Antiretroviral Therapy
His	Hydroxyethyl Starch
HIV	Human Immunodeficiency Virus
HIV-1	Human Immunodeficiency Virus Type 1
HLTR	HIV-1 LTR Target Site 1
Ig	Immunoglobulin
IgG	Immunoglobulin G
kDa	kiloDalton

KRAB	Kruppel-Associated Box Domain
LTR	Long Terminal Repeats
min	minute
mM	miliMolar
mRNA	Messenger RNA
nM	nanoMolar
NLS	Nuclear localization signal
Tat	Trans-activator of transcription
V_H	Variable Heavy Chain
VHH	Variable Domain of Camel Heavy Chain Antibody
V_L	Variable Light Chain
VNAR	Variable Domain of the Shark New Antigen Receptor
VP64	Herpes Simplex Virus Protein 64
ZF	Zinc-finger
µg	microgram
µM	microMolar

REAGENTS AND TECHNIQUES

ABTS	2,2'-Azinobis [3-ethylbenzothiazoline-6-sulfonic acid]-diammonium salt
BSA	Bovine Serum Albumin
DMEM	Dulbecco's modified Eagle´s Medium
ELISA	Enzyme-linked immunosorbent assay
HRP	Horseradish peroxidase
LB	Lysogeny Broth
PBS	Phosphate Buffer Saline

PCR	Polymerase Chain Reaction
RPMI	Roswell Park Memorial Institute Medium
SDS-PAGE	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

AMINO ACIDS

A	Alanine	L	Leucine
R	Arginine	K	Lysine
N	Asparagine	M	Methionine
D	Aspartic acid	F	Phenylalanine
C	Cysteine	P	Proline
Q	Glutamine	S	Serine
E	Glutamic acid	T	Threonine
G	Glycine	W	Tryptophan
H	Histine	Y	Tyrosine
I	Isoleucine	V	Valine

TABLE OF CONTENTS

INDEX

INDEX OF FIGURES

INDEX OF TABLES

INTRODUCTION

1.1 Antibody

1.1.1. Antibody Structure and Function

Antibody glycoproteins belong to the immunoglobulin (Ig) class, being secreted in large quantities by plasma cells. As one of the most important defense mechanism against disease, these proteins are produced by the immune system in response to foreign substances to the human body, called antigens [1]. In humans, there are five major classes of antibodies based on their ability to perform several unique binding and effector functions. They are IgA, IgD, IgE, IgG and IgM, being IgG the most abundant in serum samples and the dominant format of therapeutic antibody [2,3].

Human IgG is a heterodimer of 150 kDa formed by a constant fragment of 55 kDa (Fc - Fragment crystallizable) and two antigen-binding fragments of 45 kDa (Fab - Fragment antigen binding) linked by a flexible polypeptide called the hinge region (Figure 1) [1,3–5].

The Fab contains the variable region (Fv) which determines the specificity and affinity to the antigen. On the other hand, the Fc region recruits immune effector functions through interactions with the C1q complex of complement and a variety of neonatal Fc receptors, including FcRn [1].

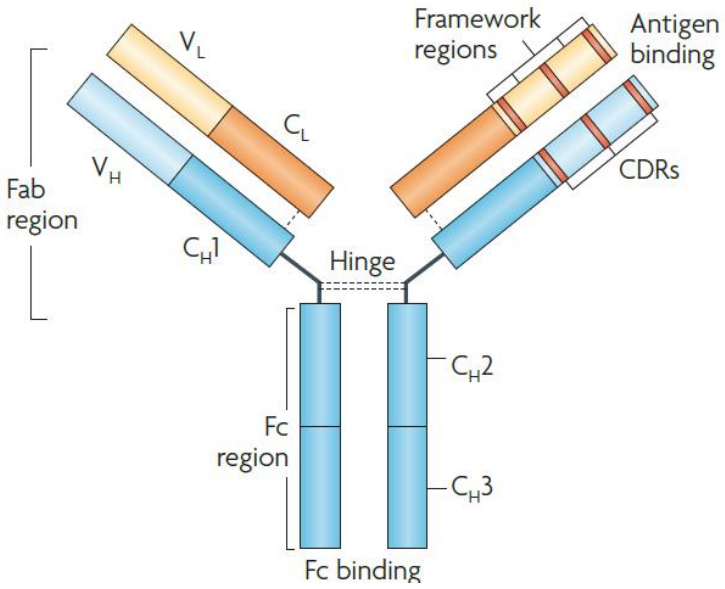


Figure 1 – Schematic representation of the structure of a conventional IgG antibody. Adapted from Hansel, T.T. Nature Reviews Drug Discovery 9, 325-338 (April 2010) [6].

IgG antibodies are 'Y'-shaped molecules comprising two identical heavy chains (H) and two identical light chains (L) linked by disulfide bonds. In the heavy chain, four globular domains define a variable region (V_H – amino-terminal) and three constant regions (C_H1, C_H2 and to C_H3). By contrast, each light chains consists in a constant region (C_L) and a variable region (V_L – amino-terminal) [4].

The variable domains of light and heavy chains form the antigen binding region. Each variable domain is composed by three loops called complementarity-determining regions (CDR1, CDR2 and CDR3), which exhibit a high variability. This variability lead to the diversity and specificity of the antibody binding, becoming it unique. The CDRs are in turn located between four conserved segments called framework regions (FRs), which are responsible for maintaining the structure of variable regions [4].

On the other hand, Fc corresponds to the segments C_H3 and C_H2 of both heavy chains. This region does not present antigen-binding activity but has a capacity to recruit effector functions, activating the immune response [4].

Often in antibody-based therapy, the immune response is essential to remove or neutralize a pathogenic infection or the pathogenic agent itself. This is because, in some cases, antibody binding to the target or its receptor cannot block the activity of the pathogen [4,7]. Consequently, the immune response is activated by effector functions of antibodies through two mechanisms of action: activation of complement-dependent cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC) (Figure 2) [3,8].

As one of the first mediators of the immune response to pathogens, CDC is a complex proteolytic cascade composed by a group of soluble proteins present in the serum, which lyse foreign cells leading to their destruction or phagocytosis induction [9,10]. The pathway is triggered by binding of the C1 complex, a serine protease, to the Fc domain of antibody bound on the surface of the target cells, inducing to the formation of the membrane attack complex (MAC) and release of powerful opsonins and anaphylatoxins [9,11].

On the other hand, ADCC can be activated by the Fc domain of antibody through interactions with FcγRs (FcγRIIIa in humans) on effector immune cells like natural killer cells (NK) [12], which lead to the destruction of the target cells by release of granzymes and perforin from NK cells [8].

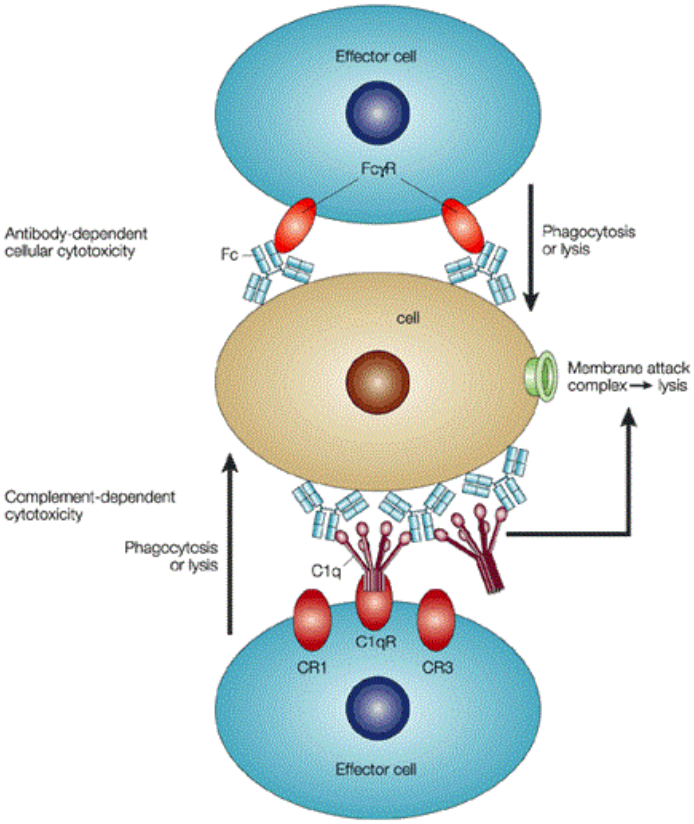


Figure 2 – The mechanisms of effector functions derived from fragment crystallizable (Fc) domain. Adapted from Carter, P. Nature Reviews Cancer 1, 118-129 (November 2001) [13].

1.1.2. Development and Production of Therapeutic Antibodies

Antigen-antibody interaction induces the proliferation and differentiation of a clonal pool of B cells, instead of just one clone with affinity for one singular epitope. As a result arises a polyclonal antibody mixture which is a mixture of antibodies against different epitopes of the same antigen. For a living organism this is an advantage, but for therapeutic or research purposes it is more advantageous to use monoclonal antibodies specific to only one epitope [4,14].

Until 1975, purification of monoclonal antibody from a polyclonal serum was not possible, when Kohler and Milstein developed a method of production of monoclonal antibodies [15]. They produced a hybrid cell with the ability of antibody production in large scale. Also known as hybridoma, it was generated fusing a B cell from an immunized donor with an immortalized myeloma cell [16].

Despite the benefits of this technology, when used in humans B cells from murine led to immunologic effects with human anti-mouse antibodies (HAMA response) [7]. This and other problems such as low circulation half-life and inability to activate human

effector functions also limited the use of this technology for therapeutic strategies. In addition, hybridoma technology has an excessively difficult and slow production and a high cost [7,17].

To try to overcome the limitations of this technology, new approaches of antibody production started to appear.

1.1.3. Antibody Engineering

The emergence of new engineered antibodies for research and therapeutic purposes was possible taking into account the existing knowledge of the structure and functions of antibodies. Genetic engineering resulted in antibody fragmentation, this is, antibodies began to be divided into smaller fragments, surpassing size limitations of IgGs. Some goals were also achieved such as lower immunogenicity, which in some cases led to longer circulation half-life and better biodistribution [18].

Thus, smaller antibody molecules like Fab (fragment antigen binding), scFv (single-chain variable fragment) or dAbs (single domain antibody) began to be produced (Figure 3) [5,18,19].

Fab consists of two variable domains and two constant domains of the light and heavy chains (V_H-C_{H1} and V_L-C_L) linked by disulfide bridges, while scFv consists of the V_L and V_H regions joined by a flexible polypeptide linker normally with 15 amino acids long (GGGGSGGGGSGGGGS). This linker increases scFv efficiency of folding and expression in *E. coli* systems [20–22]. These antibody fragments lose usually avidity, which leads to a decrease in antigen binding. However through engineering of multivalent antibody fragments this limitation can be overcome [23]. In fact, several strategies were developed in order to design multimeric scFvs which led to the formation of bivalent (diabody) and trivalent molecules (triabody) [24,25]. Such fusions can occur between equal or different scFv, in last case creating bi-specific diabodies.

The smallest functional antibody fragment, dAb, is only constituted by the variable domain of an antibody heavy chain (V_H) or light chain (V_L), thus only presenting three of the six naturally occurring CDRs of each Fab [26]. However, it seems sufficient to confer antigen binding specificity and high affinity [27,28]. The small size of dAbs results in lower immunogenicity and higher tissue penetration, similarly to the scFvs [18,29].

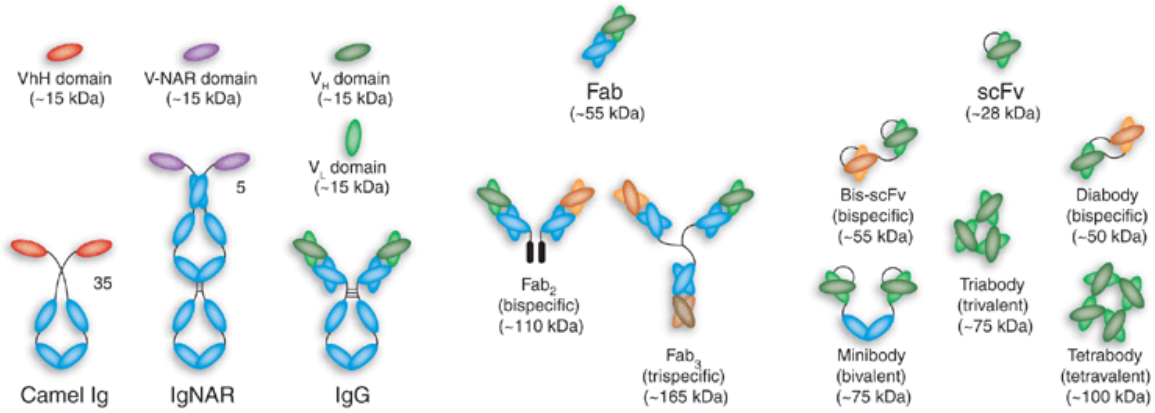


Figure 3 - Schematic representation of an IgG and different antibody fragments. Adapted from Holliger, P. Nature Biotechnology 23, 1126-1136 (September 2005) [18].

In the last few years, important developments have been made in the design, selection, and production of new antibodies. The choice of antibody format depends on the desired therapeutic purpose.

1.1.3.1. Single Domain Antibodies – VHH

Most naturally occurring antibodies are composed of two light chains folded into two domains and two heavy chains folded in four domains. Besides these conventional antibodies, camelids (including camels and llamas) and sharks also produce antibodies comprising only a single variable domain of heavy chain designated VHH or VNAR, respectively [27,28]. As it has been stated before, this recombinant antibody format is easily obtained through recombinant antibody technology [30].

The VHH domain is the smallest antibody format (12-15 kDa), which retains its antigen binding activity. It has several favorable characteristics, including high solubility, physical stability and refolding capacity [18,31]. In addition, the small size of VHHs results to relatively good tissue penetration and the ability to bind to cryptic antigenic sites [18,31,32]. For strategies against rapidly mutating viruses, e.g. HIV, this characteristic could be important for targeting their surfaces regions that are important for the viral replication cycle [33].

Conventional V_H domains of a human antibody as well as VHHs have four framework regions that form the core structure of the IgG and three CDRs that are involved in antigen binding [33]. Although in framework 2 region (positions 37, 44, 45 and

47) occur important amino acid differences. In VHHs this region is hydrophilic, becoming soluble in aqueous solution. In contrast, in V_H region is hydrophobic due to its association with cognate V_L domain [33,34].

Moreover, the CDRs of VHHs contain some specific characteristic. The CDR1 and CDR2 of V_H present a canonical structure, which changes with length of these loops and the presence of specific residues at key positions. In contrast, CDR1 and CDR2 of VHHs deviate significantly from the canonical loop structure. The CDR3 of VHHs has the capacity to form long fingerlike extensions, which can explain the binding into antigen cavities. This region is, on average, much longer than that of conventional V_H domains and is often stabilized by an interloop disulfide bond connecting the CDR3 to CDR1 or CDR2 loop [33,34].

Due to their characteristics, VHHs have many advantages for biotechnological applications, ranging from simple research tools as diagnostic reagents [18]. VHHs can be cloned into various formats due to their flexible linker, which allows simultaneous binding of multivalent antigens. The design of VHHs increases their utility for new therapeutic purposes [18,35]. Regarding production issues, dAbs like VHHs as well as conventional antibodies are expressed in mammalian cell systems. However, only dAbs are well expressed in bacterial and yeast systems [36,37].

In resume, with the recent evolution of antibody engineering, antibodies have been overcome their therapeutic limitations. In fact, the knowledge about structure and function of immunoglobulins allowed the reduction in size, reconstruction into multivalent molecules and conjugation with other proteins or peptides for the treatment of a variety of diseases.

1.2. Gene Therapy

At the end of 20th century, gene therapy emerged as a powerful alternative for the treatment or prevention of a variety of human diseases, probably due to increasing our knowledge of the structure and function of the human genome. Although, there are still few therapeutic strategies for gene manipulation.

Essentially the aim of this therapy is to correct abnormality phenotype or genotype as well as to provide cells with new functions. For this purpose, new genetic instruction are introduced into tissue of patients [38]. It is important to note that the manipulation of gene expression occurs only in somatic cells that translates into genetic correction for the patient but not for the next generations [39].

In practice this is a complex operation, due to several obstacles that must be overcome. The choice of an efficient, specific and non-toxic gene delivery system is the key point to the clinical success of gene therapy.

The most common strategies that have been developed to manipulate gene expression include recombinant viral vectors and various forms of non-viral plasmid DNA [38,40].

As mentioned above, one of the successful gene therapy systems are the viral vectors. Due to their capacity to penetrate into the cell nucleus of the host cells and exploit the cellular machinery, viral particles has been manipulated to express therapeutic genes or to infect and replicate specifically in target cells. Distinct viruses are adapted as vectors, but the most advanced are retrovirus, adenovirus and adeno-associated virus. In spite of their high transfection efficiency *in vivo*, the high immunogenicity, the limited size of transgene, the toxin production, the transient transduction and difficulty manufacture have restricted the use of these vectors in gene therapy [38,40–42].

Regarding to non-viral strategies, they include naked DNA, DNA complexed with cationic lipids and DNA condensed in particles with cationic polymers [38,43]. Although these vectors show less efficient transfection (especially *in vivo*), they represent an importantly alternative to viral vectors due to their low host immunogenicity and easily manufacture. In addition, they have no insert-size limitation [40].

Despite the clear advantages of viral and non-viral vectors, there is still a need to improve the technologies for gene manipulation, leading this therapy into benefits for medicine and biotechnology.

1.2.1. Zinc-Finger Proteins: Structure and Function

To overcome the limitations of viral delivery in gene manipulation zinc-finger proteins emerged as powerful tools for gene therapy [44,45].

Zinc-finger proteins belong to the class of DNA-binding proteins, due to their ability to bind DNA in a specific manner. A zinc-finger consists of a ββα fold coordinated by a zinc ion (Figure 4). DNA-binding residues are localized at positions -1, 3 and 6 on the surface of the α-helix, making specific contact with 3-4 nucleotides of DNA [44,45]. For this reason, critical amino acids of an zinc-finger can be designed to enable the altered protein domain to recognize a wide range of DNA sequences [44]. Therefore these characteristics of zinc-fingers make them a versatile tools for gene engineering [46].

Additionally, due to their net positive charge, zinc-fingers proteins are cell permeable which explains their innate ability to cross cell membranes. This capacity of zinc-fingers promotes direct protein delivery into various mammalian cell types [44,47].

The first application of zinc-finger technology was the creation of artificial transcription factors which modulate gene expression [44]. In fact, transcription factors are involved in the initiation and regulation of the transcription of genes. Their DNA-binding domains have the ability to bind to specific sequences of DNA also known as promoters. Therefore, transcription factors can stimulate or repress transcription of a gene of interest [48].

Indeed, when fused to a transcriptional activator, such as the VP64 transactivation domain, or a repressor, such as the Krüppel-associated box domain (KRAB), artificial zinc-fingers acquire the capacity of up- or down-regulation of a gene of interest by unknown mechanisms [44,49]. Moreover, appending zinc-fingers with other effector domains, these proteins fulfill several functions for a variety of purposes, such as integration and excision (zinc-finger recombinase), cleavage (zinc-finger nuclease), methylation (zinc-finger methyltransferase) and chromatin remodeling (zinc-finger histone methyltransferase) [44].

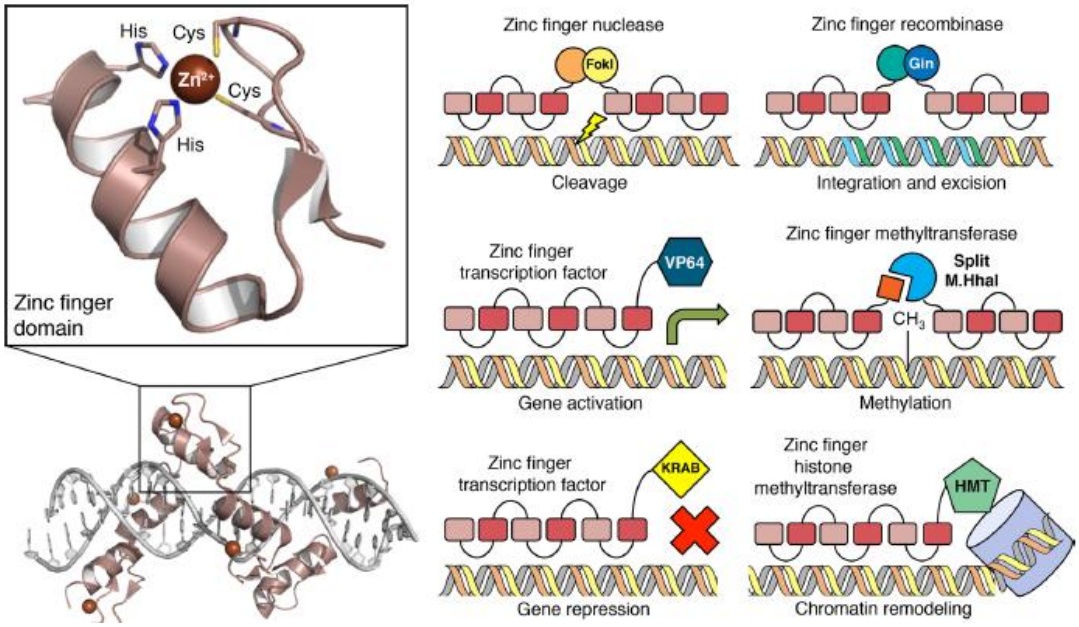


Figure 4 –Schematic representation of the structure of zinc-finger protein and the different applications of zinc-finger technology. The designed six-finger zinc-finger protein in complex with target DNA. Adapted from Gersbach, C. A. Accounts of Chemical Research 47, 2309 – 18 (August 2014) [44].

1.3. Human Immunodeficiency Virus

In 1981, the syndrome we now know as acquired immunodeficiency syndrome (AIDS) was described for the first time. It is characterized by a severe trauma to the patient immune system making him susceptible to opportunistic fatal diseases. Some years later, in 1983, the human immunodeficiency virus (HIV) was identified as the causative agent of this disease [50–52].

AIDS represents a major global public health issue, having contracted the disease more than 60 million people worldwide since its discovery, a third of whom died subsequently (WHO Global Health Observatory on AIDS). Until now two types of HIV were identified, type 1 (HIV-1) and type 2 (HIV-2). HIV-1 is more pathogenic and has higher transmission rates than HIV-2, which explains why HIV-1 is responsible for a worldwide pandemic and HIV-2 accounts for more localized epidemics, particularly in West Africa [53,54].

HIV-1 infects CD4 positive cells such as T-helper lymphocytes and macrophages [55,56], by fusing with their membranes which cause their death in less than two days [57,58]. Moreover, HIV-1 has also capacity to infect dendritic cells, natural killer-T cells and hematopoietic progenitor cells. Not only CD4 cell surface receptor enable the binding and entrance of the virus into target cells but also CC-chemokine receptor 5 (CCR5) and CXC-chemokine receptor 4 (CXCR4) determine the cellular tropism of HIV-1 [59].

To repress the replication of the HIV-1 as well as eradicate it, several therapeutic strategies have been developed. Although powerful, highly active antiretroviral therapy (HAART), a cocktail combination of three or more antiretroviral drugs that target different essential steps of the HIV-1 replication cycle [60] is still not capable to cure AIDS. For this reason, new strategies to suppress infection and to stop worldwide HIV pandemic are needed.

1.3.1. HIV-1 Genome and Structure

HIV-1 belongs to the *Lentivirus* genus from the *Retroviridae* family [61]. The retroviral genome is composed of two positive single strain RNA molecules of about 9 kb, flanked by 5' and 3' long-terminal repeats (LTRs). 5'LTR contains the viral promoter. The HIV-1 genome encodes nine open reading frames (ORF): essential genes (*gag*, *pol* and *env*) and accessory genes (*tat*, *rev*, *nef*, *vif*, *vpr* and *vpu*) (Figure 5) [62].

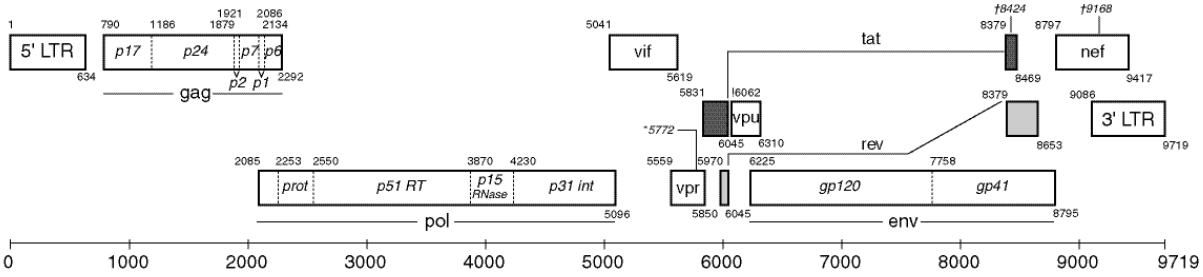


Figure 5 – Organization of the HIV-1 genome. [63]

The *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) genes encode for polyproteins which are proteolysed into individual proteins. The four Gag proteins are structural components - matrix (MA), capsid (CA), nucleocapsid (NC) and p6. The two Env proteins are present in the outer membrane of the virion and allow the binding and entrance into host cells - gp120 (surface or SU) and gp41 (transmembrane or TM). The three Pol proteins provide essential enzymatic functions and are encapsulated within the particle - protease (PR), reverse transcriptase (RT) and integrase (IN) [61,64,65].

Three of accessory proteins, Vif, Vpr and Nef are encapsulated in viral particles (Figure 6) and are responsible for modulating cellular events. Two other accessory proteins, Tat and Rev, are essential to gene regulation. The last protein, Vpu, is responsible for indirectly assisting in assembly of the virion [62,65].

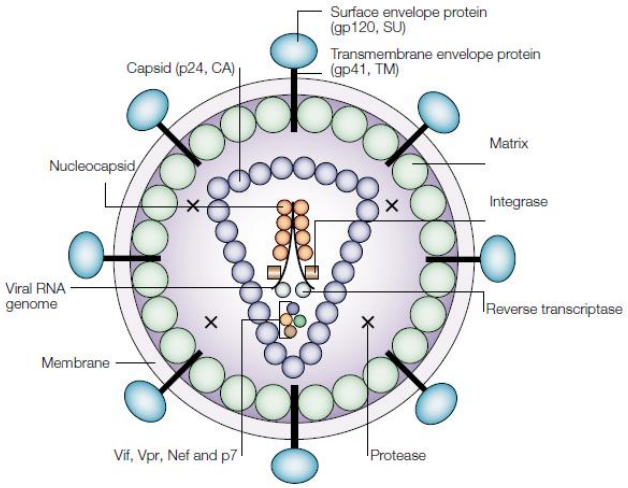


Figure 6 – HIV-1 virion. Adapted from Robinson, H. L. Nature Reviews Immunology 2, 239-250 (April 2002) [66].

1.3.2. HIV-1 Replication Cycle

HIV-1 replication events are divided into early and late phases (Figure 7). The first step in the early phase is attachment of the viral envelope glycoprotein gp120 spike to the cell surface protein CD4 [61,67]. The gp120-CD4 interaction promote exposure of a group of chemokine receptors, particularly CXCR4 or CCR5, which determine the cell tropism of HIV-1, as mentioned above [68]. Subsequently, TM gp41 undergoes conformational changes in order to expose a "fusion peptide" that triggers the membrane fusion process and the entry of the viral capsid into the cytoplasm of the host cell [65,67,69].

Once inside the cell, the viral core suffers an "uncoating" process that involves the dissociation of the capsid, which is essential for the progress of reverse transcription [67]. Following uncoating, reverse transcriptase copy the viral RNA genome into a double-stranded linear DNA, which is integrated into the host genome with the help of the integrase, completing the early phase [61,65,70,71].

Following successful integration, the provirus is transcribed by the host RNA polymerase II into spliced and unspliced mRNA transcripts [61]. Initially, only short spliced mRNAs are translated and these code for proteins Tat, Rev and Nef. Tat is an essential transcriptase activator that binds to the trans-activating response element (TAR), located at the 5'LTR promoter, increasing the rate of HIV-1 transcription and elongation. Rev is responsible for the RNA export of single-spliced (*env*, *vif*, *vpu* and *vpr*) and unspliced transcripts (*gag* and *gagpol*) from the nucleus to the cytoplasm [65]. Once in the cytoplasm, Gag, Gagpol and Env proteins gp120 and gp41 are translated. Afterwards, proteins are transported to the plasma membrane through vesicular, cytoskeletal and other routes.

These proteins are encapsulated and immature virions are released from the cell. Consequently, the viral protease triggers the maturation of virions with a dramatic reorganization of their core and the acquisition of infectious capacity [61].

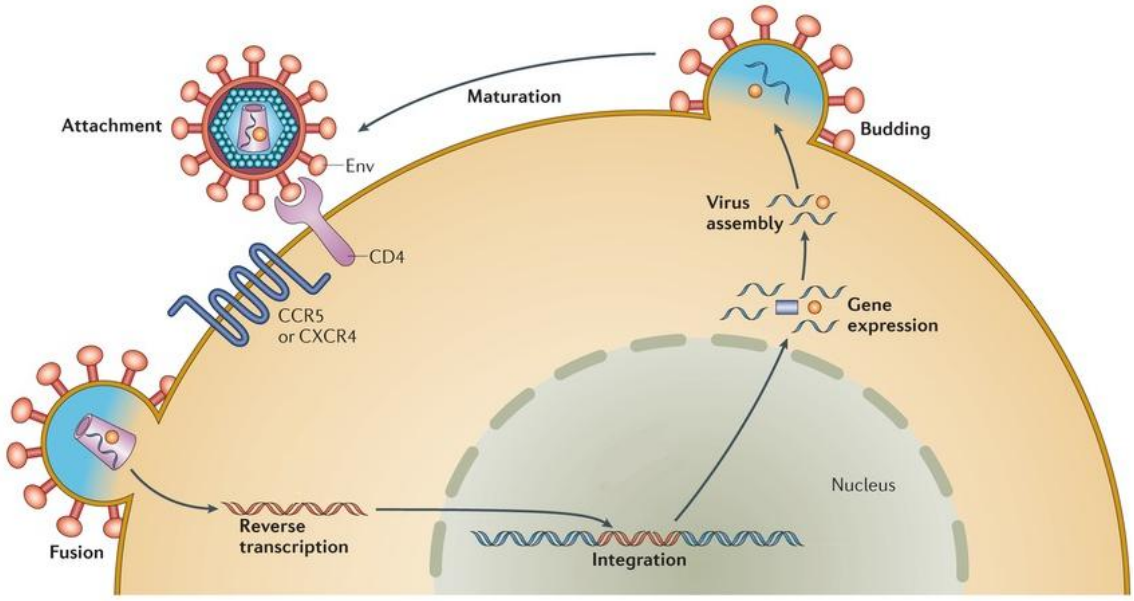


Figure 7 – HIV-1 replication cycle. Adapted from Laskey, S. B. Nature Reviews Microbiology 12, 772-780 (September 2014) [72].

1.3.3. Gene Manipulation Strategies against HIV-1

To reduce morbidity and mortality of HIV infected individuals, several strategies have been developed to date. Although, current anti-retroviral treatment is still not capable to eradicate the virus. Moreover, the emergence of drug resistance and drug toxicity often limited the treatment.

With this in mind, the development of new therapeutic strategies against HIV is urgently required. Gene manipulation strategies have demonstrated a great potential to treat this disease. To inhibit HIV replication, anti-HIV gene therapies consist in transgenes delivery into HIV-susceptible cells, turning them resistant to infection or immune against HIV-1 antigens [73].

Other therapeutic strategies based on gene targeting have been developed, such as zinc-finger-based therapies, due to their capacity to bind DNA in a specific and efficient manner [45]. Although, the small HIV-1 genome size and their high mutation rate limit the potential sites for engineered zinc-fingers. The LTR promoter is an attractive target in the HIV-1 genome since it is one of the few accessible conserved regions [74].

In two independent studies [75,76] zinc-finger proteins fused to repression domains have been designed to target the HIV-1 LTR promoter and drastically reduced HIV-1 replication. Repression was observed with no significant cytotoxicity. Afterwards, Kim *et*

al. extended these studies, creating a zinc-finger protein linked to two repressor domains designed to target Sp1-site binder. POZ of FBI-1 specifically blocks the DNA binding domain of Sp1 and TatdMt, a mutant of Tat. In fact, this artificial zinc-finger inhibited HIV-1 replication [77]. In 2005, Scott R. Eberhardy *et al*. designed several zinc-finger proteins linked to repressor domains designed to bind at the primer-binding site (PBS) and repress transcription from the HIV-1 LTR. PBS is the most highly conserved region in the HIV-1 LTR promoter. One of these proteins showed the ability to inhibit viral replication over the course of several weeks (90% inhibition), and no significant cytotoxicity was observed [78].

In resume, these studies suggest that zinc-finger-based therapies have the potential to be applied in the clinical, due to the innate ability of zinc-fingers to cross cell membranes and specifically bind to DNA. However, in order to try erradicate the virus and to stop worldwide HIV-1 pandemic, the limitations of this strategy must be overcome such as the specificity of zinc-fingers to the target cells. Moreover, the development of new therapeutic strategies are still required.

1.4. Aims

The increasing knowledge of gene regulation led to the emergence of gene therapy as a powerfull tool for the treatment and prevention of a variety of human diseases. New strategies for gene manipulation have been developed to date, such as zinc-finger proteins. Due to their capacity to bind DNA in a specific and efficient manner, zinc-fingers possess a great potential for gene targeting and regulation. However this strategy presents several limitations that must be overcome in the future, in particular their low specificity to the target cells. With this in mind, new zinc-finger delivery strategies are needed to improve their specificity to the target cells. Antibody fragments have demonstrated to have high specificity which make them a powerful tools to overcome this limitation presented by zinc-finger proteins.

Within this context the aim of this research project is to optimize the therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers previously developed in our laboratory. To validate our strategy and as a proof-of-concept, we choose HIV-1 as the disease model since Acquired Immune Deficiency Syndrome (AIDS) is a disease which represents a major global public health issue.

In order to compare to the previously studied protein (Cunha-Santos *et al*., unpublished results), we will engineer and construct four alternative bispecific proteins of an artificial zinc-finger with the KRAB repressor domain designed to repress the transcription from the HIV-1 LTR promoter (KRAB-HLTR3) coupled to a CXCR4-specific nanobody (VHH). In two of the constructions, the cathepsin B cleavage site will be introduced between the KRAB-HLTR3 and the anti-CXCR4 VHH, in order to facilitate the release of the zinc-finger in the target cells. For the same purpose, in one of the constructions, MMMP-9 cleavage site will be introduced in the same position as described above.

Thus, the main goals of the present work are the following:

1) Construction of a zinc-finger protein fused with the anti-CXCR4 VHH;

2) Optimization of the expression and purification conditions of the bispecific proteins;

3) Characterization of the bispecific proteins binding activities against HIV-1 LTR promoter and against CXCR4 receptor;

4) Evaluation of bispecific proteins ability to repress transcription from the HIV-1 LTR promoter through reporter gene assays.

MATERIALS AND METHODS

2.1 Cloning of recombinant proteins

2.1.1. KRAB-HLTR3 (KH)

A fragment encoding HLTR3 artificial zinc-finger fused to the Krüppel-associated box repressor domain (KRAB) [75] was cloned into the bacterial expression vector pET-21a (+). KRAB-HLTR3 was previously described to repress transcription from the HIV-1 long terminal repeat (LTR), and it was kindly provided by C. Cunha-Santos (João Gonçalves Lab, unpublished results).

2.1.2. Anti-CXCR4 VHH-KRAB-HLTR3 Cathepsin B (CKH-Cat)

The protein gene was synthesized by Invitrogen™ fusing DNAs encoding anti-CXCR4 VHH clone [31] and the KRAB-HLTR3 zinc-finger [75], adding sequences encoding peptide tags for purification (His8) and detection (HA) on the C-terminal. Cathepsin B cleavage site was introduced between the anti-CXCR4 VHH and the KRAB-HLTR3. The introduced sequence is 5'-CCC CTG AAG CCC GCC AAG AGC GCC AGA AGC-3'.

A fragment encoding anti-CXCR4 VHH-KRAB-HLTR3-Cathepsin B was digested with NdeI and NcoI restriction enzymes and subcloned into the bacterial expression vector pET-21a (+).

2.1.3. Anti-CXCR4 VHH-KRAB-HLTR3 (CKH)

The pET-21a(+)-CKH-Cat was digested with the EcoRI restriction enzyme, deleting the cathepsin B cleavage site. The resulting fragment was gel purified (Zymo Research, USA) and we performed a religation of the vector.

2.1.4. Anti-CXCR4 VHH-KRAB-HLTR3 MMP9 (CKH-MMP9)

A fragment encoding the anti-CXCR4 VHH clone [31] was amplified by PCR with primers 1 and 2, adding the matrix metalloproteinase 9 (MMP-9) cleavage site at the fragment 3´end. PCR fragment was gel purified, digested with the NheI and SacI restriction enzymes and cloned in pET-21a (+)-CKH NheI and SacI digested, substituting the anti-CXCR4 VHH. The introduced sequence of MMP-9 cleavage site is 5'-AAA ATA CCA AGA ACG TTG ACG -3'.

2.1.5. KRAB-HLTR3-Anti-CXCR4 VHH (KHC)

A fragment encoding anti-CXCR4 VHH clone [31] was fused to KRAB-HLTR3 zinc-finger [75] and cloned into the bacterial expression vector pET-21a (+). KHC was kindly provided by C. Cunha-Santos (João Gonçalves Lab, unpublished results).

2.1.6. KRAB-HLTR3-Anti-CXCR4 VHH Cathepsin B (KHC-Cat)

A fragment encoding the anti-CXCR4 VHH clone [31] was amplified by PCR with primers 3 and 4, adding cathepsin B cleavage site at the fragment 5´end. PCR fragment was gel purified, digested with the SpeI and NcoI restriction enzymes and cloned in pET-21a (+)-KHC SpeI and NcoI digested, substituting the anti-CXCR4 VHH. The introduced sequence of cathepsin B cleavage site is 5'-CCC CTG AAG CCC GCC AAG AGC GCC AGA AGC-3'.

All constructions were carried out by DNA digestion with enzymes from Thermo Fisher Scientific (UK). T4 DNA Ligase (Thermo Fisher Scientific, UK) was used in vector-insert ligations. Polymerase Chain Reaction (PCR) was performed in Doppio thermocycler (VWR International, USA) using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, UK). All primers used for PCR reactions are presented in table 2 in annexes. PCR conditions program used is presented in table 3 in annexes. Clones were screened by digestion with appropriate restriction enzymes and resolved by agarose gel electrophoresis. Positive clones sequence was confirmed by DNA standard sequencing (GATC Biotech, Germany).

2.2. Expression and purification of proteins

2.2.1. CKH-Cat

The plasmid construction was expressed in *E.coli* strain BL21 (DE3). 500 mL of LB, containing 100 µg/ml ampicillin (NZYtech, Portugal) was inoculated with 5 ml of overnight culture of bacterial cells and grown to exponential phase (A_{600} = 0.6-0.9) at 37 °C. Expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Thermo Fisher Scientific, UK) and growth during 4 h at 37 °C. Cells were harvested by centrifugation (10,000 x g for 15 min at 4 °C) and resuspended in 20 ml buffer A (50 mM HEPES, 1 M NaCl, 10 mM imidazole, 5 mM

$CaCl_2$, 2 M urea, 1 mM β-mercaptoethanol (pH 8)), and lysed by sonication for 30 min. Cell pellet/insoluble fraction was collected by centrifugation (12,000 × g for 30 min at 4 °C), resuspended in 10 ml buffer A (50 mM HEPES, 1 M NaCl, 10 mM imidazole, 5 mM $CaCl_2$, 2 M urea, 1 mM β-mercaptoethanol (pH 7.4)), and lysed by sonication for 30 min. Cell pellet/insoluble fraction was recollected by centrifugation (12,000 × g for 30 min at 4 °C) and resuspended in 25 ml buffer B (50 mM HEPES, 1 M NaCl, 10 mM imidazole, 5 mM $CaCl_2$, 6 M urea, 1 mM β-mercaptoethanol (pH 7.4)). The protein pellet resuspended in buffer B is subjected to a denaturation/solubilization step overnight at 4 °C in a vertical rotator (Stuart rotator, Dynalab). The solubilized protein (in buffer B) was submitted to centrifugation (12,000 × g for 60 min at 4 °C) for removal of remaining cell debris and insolubilized protein. The solubilized protein solution was filtered through a 0.45 µm syringe filter (Sarstedt, Germany).

The protein was subsequently applied to a Ni-NTA His GraviTrap™ Column (GE Healthcare, UK) and purified according to the manufacturer's instructions for purification under denaturing conditions. The protein was eluted in 3 ml buffer C (50 mM HEPES, 1 M NaCl, 500 mM imidazole, 5 mM $CaCl_2$, 6 M urea, 1 mM β-mercaptoethanol (pH 7.4)).

The recombinant protein refolding/buffer exchange was performed in HiTrap™ Desalting Column (GE Healthcare, UK) with Peristaltic Pump P-1 (GE Healthcare, UK) according to the manufacturer's instructions with Zinc Buffer A (20 mM HEPES, 1 mM $MgCl_2$, 100 µM $ZnCl_2$, 90mM KCl (pH 7.4)) as a final buffer. After purification, the protein was characterized and protein purity was analyzed by SDS-PAGE, as well as protein concentration was quantified by Bradford method. This experimental protocol was adapted from Cunha-Santos *et al.* [79].

2.2.2. CKH

The protocol used for the expression and purification of the protein is identical to the protocol described in 2.2.1.

2.2.3. KHC

The protocol used for the expression and purification of the protein is identical to the protocol described in 2.2.1.

2.2.4. KHC-Cat

The protocol used for the expression and purification of the protein is identical to the protocol described in 2.2.1.

2.2.5. KH

The protocol used for the expression and purification of the protein is identical to the protocol described in 2.2.1.

2.3. Coomassie staining

Protein separation was performed according to the method of Laemmli in 15 % polyacrylamide gels (SDS-PAGE). Following electrophoresis, the gel was placed in staining solution (40 % methanol, 10 % acetic acid, 0,025 % Coomassie Brilliant Blue). The gel was incubated for 1 h to overnight in the staining solution. The gel was destained with several changes of destain solution (30 % methanol, 10 % acetic acid) until the background is transparent. All staining destaining steps were done on a rotary shaker with gentle mixing.

2.4. Western blot

Protein separation was performed according to the method of Laemmli in 15 % polyacrylamide gels (SDS-PAGE). Once separated, the proteins were electrotransferred into a nitrocellulose membrane (GE Healthcare, UK). Membrane was blocked with a 5% milk-TBS 0.1 % Tween20 solution for 1 h and proteins were detected using a HRP-conjugated HA-tag antibody (Roche, Germany) diluted 1:5000 in 5 % milk-TBS 0.1 % Tween20 solution for 1 h at RT with agitation. Membrane was washed 5 times with TBS 0.1 % Tween20 solution. Antibody detection was made with Immobilon™ Western Chemiluminescent HRP substrate (Millipore, USA). Membrane was incubated with HRP substrate for 5 min at RT and then revealed in a chemiluminescence film Amersham Hyperfilm™ ECL (GE Healthcare, UK).

2.5. ELISA to evaluate the bispecific proteins binding ability to HLTR3 binding site

Binding properties of all recombinant proteins were determinate in 96 well flat bottom, high binding non-sterile, polystyrene ELISA plates (Corning, USA) coated with Streptavidin (Thermo Fisher Scientific, UK) in PBS (400 ng/well) overnight at 4 °C. The plate was washed with dH_2O and incubated for 1 h at 37 °C with a specific sequence of oligonucleotides (HLTR3 binding site [75]) in PBS (25 ng/well). After 1 h blocking with 3% bovine serum albumin (BSA) (Sigma-Aldrich, USA) in Zinc Buffer A (ZnBA), purified proteins were incubated for 2 h at room temperature (RT). Serial dilutions of purified proteins (200 pM to 200 nM) were diluted in 1 % BSA in ZnBA/3µg/well herring sperm DNA (Promega, USA). After washing 5 times with dH_2O, 25 µl/well of HRP-conjugated HA-tag antibody (Roche, Germany, 1:1000 dilution in 1 % BSA (in ZnBA)) was incubated for 1 h at 37 °C. The plate was then washed 10 times with dH_2O and developed with an HRP substrate, ABTS solution (Calbiochem, Germany) and absorbance was measured at 405/495 nm in a microplate reader (Bio-Rad, USA). GraphPad Prism Software version 5 was used for data analysis.

2.6. Cell lines and culture conditions

Human embryonic kidney (HEK) 293T (ATCC, USA) is a highly transfectable cell line derived of the 293 cell line into which the temperature sensitive gene for SV40 T-antigen was inserted. HEK 293T cell line was cultured in Dulbecco's minimal essential medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/mL Penicillin, 100 µg/mL Streptomycin, and 0.25 µg/mL Amphotericin B (DMEM-10).

Jurkat E6-1 (NIH AIDS Research and Reference Reagent Program, USA) is a human CD4+ T leukemia cell line. Jurkat cell line was cultured at a density of 1×10^6 cells/ml of RPMI-1640 medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/mL Penicillin, 100 µg/mL Streptomycin, and 0.25 µg/mL Amphotericin B (RPMI-10).

HeLa-Tat-III (NIH AIDS Research and Reference Reagent Program, USA) is a cell line which constitutively express HXBc2 tat. It was cultured in Dulbecco's minimal essential medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/mL Penicillin, 100 µg/mL Streptomycin, and 0.25 µg/mL Amphotericin B (DMEM-10).

HeLa-Tat-III/LTR/d1EGFP (NIH AIDS Research and Reference Reagent Program, USA) is a cell line derived from HeLa-Tat-III cells. It was transfected with d1EGFP under the control of HIV-1 LTR promoter, thus, the cells also constitutively express d1EGFP. This cell line was cultured in Dulbecco's minimal essential medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/mL Penicillin, 100 µg/mL Streptomycin, and 0.25 µg/mL Amphotericin B (DMEM-10).

All cell lines were cultivated in T75 tissue culture flasks (75 cm^3) (Sarstedt, Germany), at 37 °C with 5 % CO_2. Every cell culture media and reagents, otherwise indicated, were from Lonza (Switzerland).

2.7. Flow cytometry analysis

2.7.1. Flow cytometry assay to evaluate specific binding to CXCR4 receptor

For each assay condition, 2 x 10^5 Jurkat E6-1 T-cells (NIH AIDS Research and Reference Reagent Program, USA) were seeded per well in 96-well plates (Sarstedt, Germany) and were incubated with 100 nM of recombinant proteins for 1 h at 4 °C. Following 1 h of incubation, cells were washed twice with PBS and were incubated with FITC conjugated HA-tag antibody (Santa Cruz Biotechnology, USA) for 30 min at 4 °C. After cells were washed twice with PBS. Bispecific proteins binding to CXCR4 was detected by excitation at 488 nm and detection at 525 nm. Flow cytometry analysis was performed in Guava® easyCyte HT (Millipore, USA), by acquirement of 2.000-gated events from each sample. Data were analyzed using FlowJo software (Tree Star, USA).

This experimental protocol was designed and optimized by C. Cunha-Santos (João Gonçalves Laboratory, unpublished results).

2.7.2. Flow cytometry assay to evaluate the inhibition of the HIV-1 LTR promoter

For each assay condition, 2 x 10^5 HeLa-Tat-III/LTR/d1EGFP cells (NIH AIDS Research and Reference Reagent Program, USA) were seeded per well in 24-well plates (Sarstedt, Germany) and were incubated with different concentrations (10 nM, 25 nM, 50 nM, 100 nM, 150 nM) of recombinant proteins at 37 °C for 3 h. Afterwards, cells

were washed with warm PBS and detached from the plate with cell dissociation buffer (Gibco, USA). Cells were washed with cold PBS and incubated with 2 µl annexin V conjugate (Life Technologies, USA). GFP expression was detected by excitation at 488 nm and detection at 525nm. Dead cells were detected by excitation at 405 nm and detection at 452 nm with Pacific Blue Filter. Flow cytometry analysis was performed in BD LRSFortessa™ (BD Biosciences, USA), by acquirement of 5.000-gated events from each sample. Data were analyzed using FlowJo software (Tree Star, USA).

This experimental protocol was designed and optimized by C. Cunha-Santos (João Gonçalves Laboratory, unpublished results).

2.8. Cloning and expression of Trastuzumab-E2C

2.8.1. Trastuzumab-E2C

A fragment encoding the KRAB-E2C [80] was amplified by PCR with primers 5 and 6. PCR fragment was gel purified and subcloned into the mammalian expression vector pCEP4 using HindIII restriction enzyme. pCEP4 was previously cloned with a fragment encoding Trastuzumab, which was kindly provided by Dr. Christoph Rader [81,82].

2.8.2. Trastuzumab-E2C-Cathepsin B

The protocol used for the cloning of the protein is identical to the protocol described in 2.8.1. PCR was performed with primers 6 and 7. Cathepsin B cleavage site was introduced between the Trastuzumab and the KRAB-EC2. The introduced sequence is 5'-CCC CTG AAG CCC GCC AAG AGC GCC AGA AGC-3'.

2.8.3. Trastuzumab-E2C-MMP9

The protocol used for the cloning of the protein is identical to the protocol described in 2.8.1. PCR was performed with primers 6 and 8. MMP-9 cleavage site was introduced between the Trastuzumab and the KRAB-EC2. The introduced sequence is 5'-AAA ATA CCA AGA ACG TTG ACG -3'.

2.8.4. Transfections

HEK293T cells were transfected by the calcium phosphate method [83]. 5 x 10^5 cells were seeded in each well of 6-well plates (Sarstedt, Germany). 24 h after, cells were transfected with 5 µg of total DNA according to protocol. Cell medium was changed the next day and 48 hours after transfection cells were harvested.

2.8.5. Immunoprecipitation and Western blot analysis

Transfected cells were washed twice with phosphate-buffer saline (PBS) solution and lysed with a RIPA buffer (25 nM Tris-HCl pH 7.6, 150 mM NaCl, 1 % NP-40, 1 % Sodium Deoxycholate, 0,1 % SDS, H_2O), supplemented with complete EDTA-free Protease Inhibitor Cocktail Tablets (Roche, Germany) on ice. Thirty minutes after, cells were centrifuged at 14,000 x g for 30 min at 4 °C and cell supernatants were recovered.

To precipitation of the immune complexes, both supernatants (from cell transfection plates and lysates) were exposed to native Protein A Sepharose 4 Fast Flow (GE Healthcare, UK). Following 1 h of incubation at 4 °C, the complexes were centrifuged (12 000 x g, 20 seconds) and the pellets were washed 3 times with PBS and once with wash buffer (50 mM Tris, pH 8). The final pellets were suspended in sample buffer (1 % SDS, 100 mM DTT, 50 mM Tris, pH 7.5).

Total protein of each sample was resolved in a 12 % SDS-polyacrylamide gel. Proteins were electrotransferred into a nitrocellulose membrane. Membrane was blocked with a 5 % milk-TBS 0.1 % Tween20 solution for 1 h and proteins were detected using a HRP-conjugated HA-tag antibody (Roche, Germany) diluted 1:5000 in 5 % milk-TBS 0.1 % Tween20 solution for 1 h at RT with agitation. Membrane was washed 5 times with TBS 0.1 % Tween20 solution. Antibody detection was made with Immobilon™ Western Chemiluminescent HRP substrate (Millipore, USA). Membrane was incubated with HRP substrate for 5 min at RT and then revealed in a chemiluminescence film Amersham Hyperfilm™ ECL (GE Healthcare, UK).

RESULTS

3.1 Construction, expression and purification of recombinant proteins

For the development of our therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers we initially constructed several zinc-finger fusion proteins, designed to target specifically the HIV-1 LTR promoter (Figure 15 in annexes) and repress viral transcription. These constructions are variants of the previously validated fusion protein KHC as antibody delivery strategy to gene manipulation (Cunha-Santos *et al*., unpublished results).

CKH, CKH-Cat, CKH-MMP9 and KHC-Cat constructions were generated by fusing anti-CXCR4 VHH [31] to the KRAB-HLTR3 zinc-finger [75], including a histidine tag (His8) followed by a hemagglutinin tag (HA) on the C-terminal. KHC and KH (only HLTR3 fused to the KRAB repressor domain) were previously constructed the same way by C. Cunha-Santos (Joao Goncalves Laboratory, unpublished results). In KHC-Cat and CKH-Cat, cathepsin B cleavage site was introduced between the KRAB-HLTR3 and the anti-CXCR4 VHH. Cathepsin B is a lysosomal cysteine protease and their main function is the degradation of proteins that have entered the lysosomal system from outside the cell or from other compartments within the cell [84,85]. For this reason, cathepsin B cleavage site was introduced in order to facilitate the release of the zinc-finger in the target cells. On the other hand, matrix metalloproteinases (MMPs) are proteolytic enzymes presented in extracellular matrix, which are capable of degrading and processing almost all components of the extracellular matrix [86]. Therefore, in CKH-MMP9, MMMP-9 cleavage site was introduced in the same position with the purpose described above.

The bispecific proteins obtained present 1320 amino acid residues with a calculated molecular weight of ~55 kDa. KH presents 890 amino acid residues with a calculated molecular weight of ~35 kDa. Unmodified anti-CXCR4 VHH was kindly provided as purified protein (C. Cunha-Santos, João Gonçalves Lab, unpublished results) and used as a control in the next assays. This protein presents a molecular weight of ~16 kDa. All constructions are shown in figure 8.

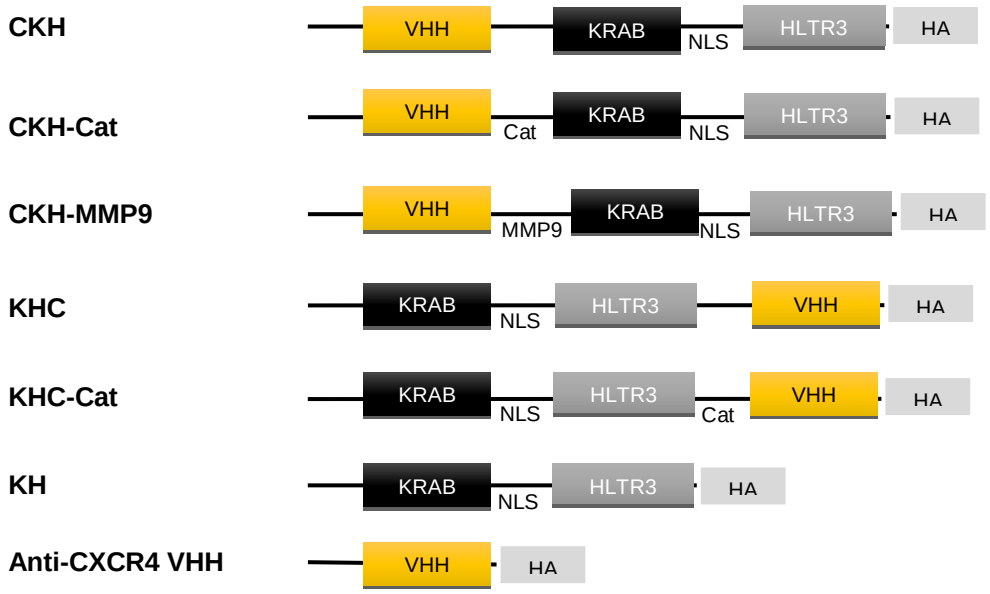


Figure 8 - Schematic representation of recombinant proteins constructions.

CKH, CKH-Cat, KHC, KHC-Cat and KH were expressed in *E. coli* BL21 (DE3) and insoluble fractions were purified by Immobilized Metal Affinity Chromatography (IMAC) and afterwards were submitted to buffer exchange (see materials and methods section). Regarding CKH-MMP9, this construct presented residual protein expression in the soluble and insoluble fractions, which indicates that this specific position on VHH framework cannot support MMP-9 cleavage site. This alteration affects the antibody stability/solubility.

SDS-PAGE and Western Blot (data not shown) results showed a single protein band with the expected molecular weights for the recombinant proteins under reducing conditions (Figure 9).

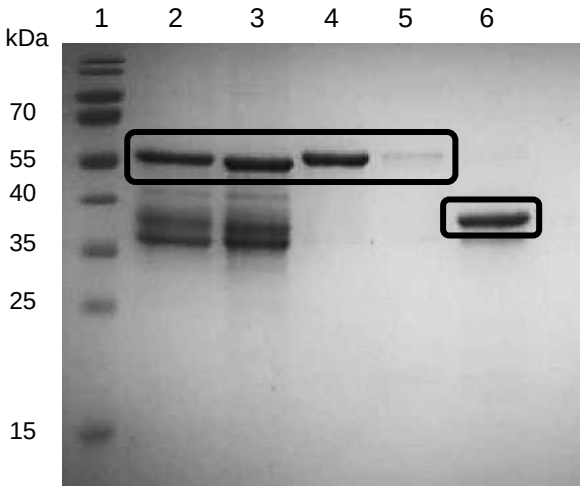


Figure 9 – SDS-PAGE analysis of purified proteins. Gel was stained with Coomassie brilliant blue. Legend: 1) Protein Ladder; 2) KHC-Cat; 3) KHC; 4) CKH-Cat; 5) CKH; 6) KH.

Purification yields for the different proteins are shown on table 1.

Table 1 – Yield of recombinant proteins per 500 mL of bacteria culture.

Construction	Total amount (µg)
CKH	126
CKH-Cat	205
KHC	309
KHC-Cat	300
KH	138

3.2. ELISA to evaluate the bispecific proteins binding ability to HLTR3 binding site

After purification it was necessary to study the functionality of proteins. Preliminary binding assays were performed by ELISA (see materials and methods section) using oligonucleotides representing HLTR3 binding site [75] as antigen.

Results shown in figure 10 demonstrate that CKH, CKH-Cat and KHC-Cat bind specifically to HLTR3 binding site in a concentration-dependent manner, similarly to KHC and KH.

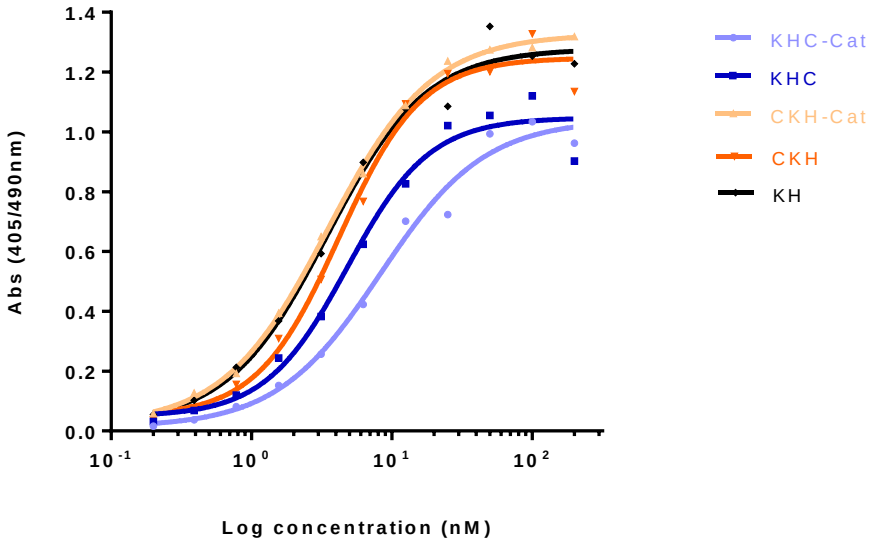


Figure 10 – ELISA assays to evaluate binding of the bispecific proteins to HLTR3 binding site. Assay was performed using 25 ng/well of HLTR3 binding site as antigen. 3 % BSA (in ZnBA) was used for blocking

and as negative control. Serial dilutions of purified proteins (200 pM to 200 nM) were diluted in 1 % BSA in ZnBA/3µg/well herring sperm DNA. Detection was achieved using an HRP-conjugated HA-tag antibody.

Importantly, binding of CKH-Cat and CKH was similar to the binding of KH, which was used as one of the positive controls. These constructions, which contain HLTR3 zinc-finger on the C-terminal, presented higher binding to HLTR3 binding site than KHC-Cat and KHC, the other positive control.

3.3. Flow cytometry assays to evaluate specific binding to CXCR4 receptor

After evaluation of proteins binding specificity to HLTR3 binding site by ELISA assay, it was necessary to verify the proteins specific binding to CXCR4 receptor. Since all constructions are bispecific proteins it was necessary to evaluate both function domains.

The chemokine receptor CXCR4 is a transmembrane receptor and consequently it is impossible to isolate it [87,88]. Since CXCR4 receptor is highly expressed in Jurkat E6-1 T-cells, proteins specific binding was performed in this cell line. In addition, CXCR4 receptor internalizes shortly after ligand attachment [89]. For this reason and in order to evaluate CXCR4 surface binding, the assays were performed at 4 °C. The proteins binding to CXCR4 receptor at the surface was detected with FITC-conjugated HA-tag antibody by flow cytometry.

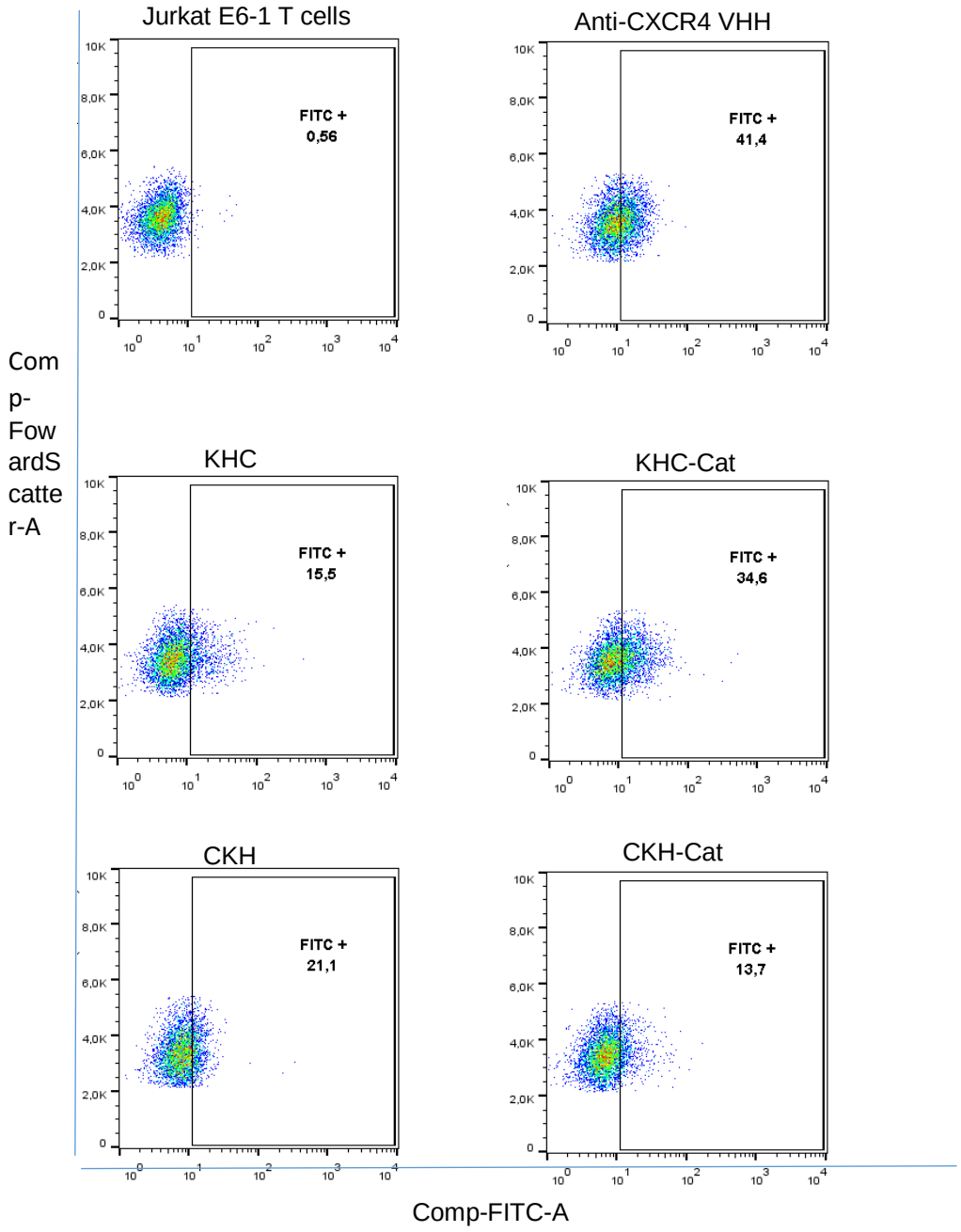


Figure 11 – Flow cytometry assay to evaluate specific binding to CXCR4. For each assay condition, 2 x 10^5 Jurkat E6-1 T-cells were seeded per well in 96-well plates and were incubated with 100 nM of recombinant proteins for 1 h at 4 °C. Cells were washed twice with PBS and incubated with FITC-conjugated HA-tag antibody for 30 min at 4 °C. Afterwards, cells were washed twice with PBS. Bispecific proteins binding to CXCR4 were detected by excitation at 488 nm and detection at 525 nm. Flow cytometry analysis was performed in Guava® easyCyte HT, by acquirement of 2.000-gated events from each sample. Values are relative to Jurkat cells treated with FITC-conjugated HA-tag antibody.

Regarding CXCR4 surface binding, as shown in figure 11, the recombinant proteins bind specifically to CXCR4 receptor like KHC. KHC-Cat presents a higher binding ability than CKH, and also higher than KHC and CKH-Cat. The actual number of FITC positive cells was 34,6 % for KHC-Cat, 21,1 % for CKH, 15,5 % for KHC and 13,7 % for CKH-Cat. Unmodified anti-CXCR4 VHH, which was used as a positive control, presents the highest binding ability, with 41,4 % of FITC positive cells.

3.4. Flow cytometry assays to evaluate the inhibition of the HIV-1 LTR promoter

The next step was to evaluate the bispecific proteins ability to gene manipulation, in particular to repress transcription from the HIV-1 LTR promoter. For this purpose, we assessed the repression of transcription of a reporter driven by the HIV-1 LTR.

The assay was performed in HeLa-Tat-III/LTR/d1EGFP cell line, a HeLa-Tat-III derivative that contains a destabilized green fluorescent protein (GFP) gene reporter [90], under the control of the LTR promoter. This cell line constitutively express GFP, presenting its fluorescence emission spectrum in green (520 nm). Since GFP has a half-life of about 1 h, cells were incubated with different concentrations of recombinant proteins for 3 h at 37 °C. As a negative control, unmodified anti-CXCR4 VHH was also tested.

When proteins bind to HIV-1 LTR promoter and repress transcription of GFP gene, the target cells decrease its green fluorescence emission, it was allows to be detected by flow cytometry (Figure 12).

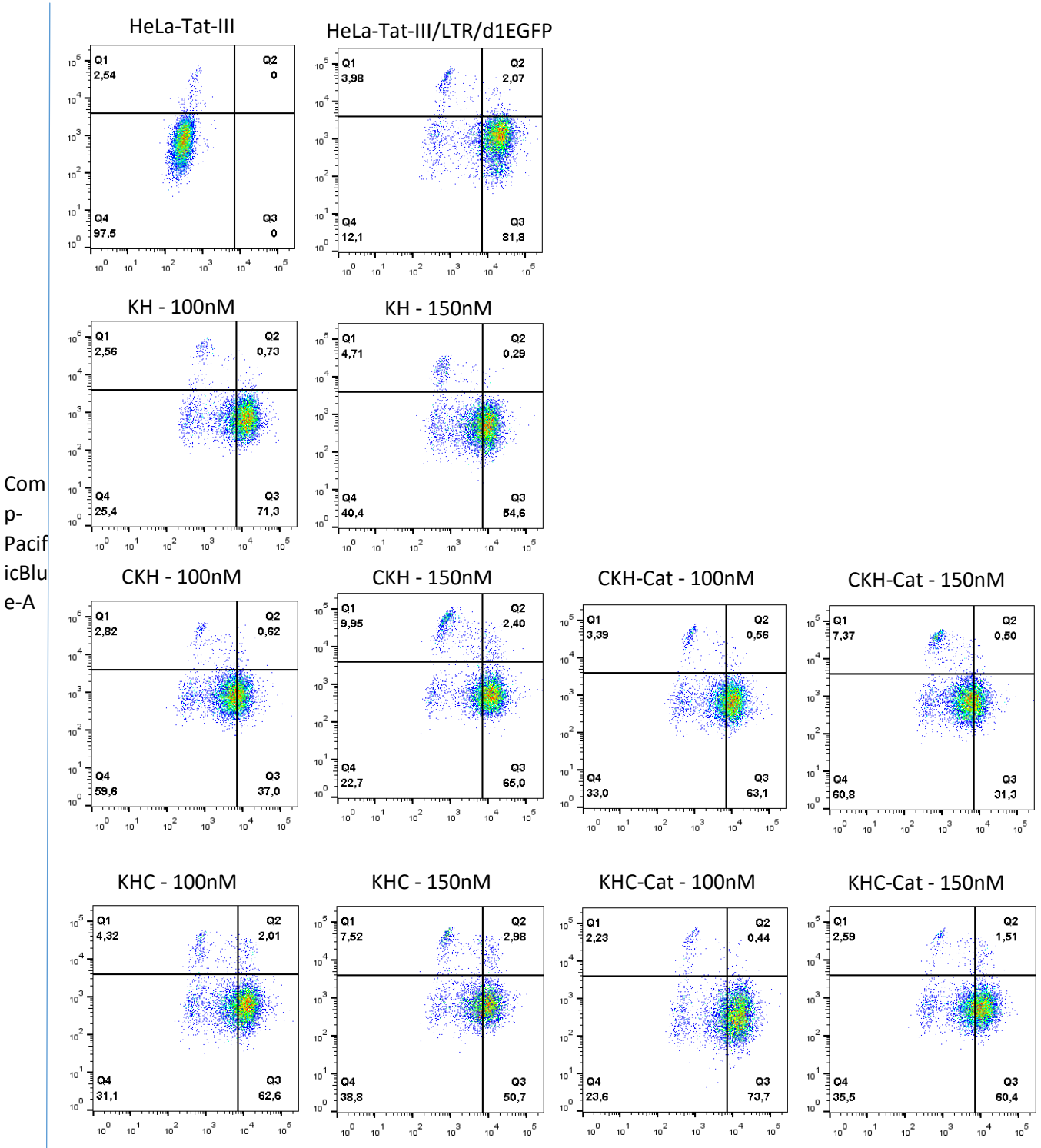


Figure 12 – Flow cytometry assay to evaluate the transcriptional repression of GFP gene driven by the HIV-1 LTR promoter. For each assay condition, 2 x 10^5 HeLa-Tat-III/LTR/d1EGFP cells were seeded

per well in 24-well plates and were incubated with different concentrations (10 nM, 25 nM, 50 nM, 100 nM and 150 nM) of recombinant proteins for 3 h at 37 °C. (Continues on the next page)

Afterwards, cells were washed with warm PBS and detached from the plate with cell dissociation buffer. Cells were washed with cold PBS and incubated with 2 µl annexin V conjugate. As a negative control, HeLa-Tat-III cell line was used. GFP expression was detected by excitation at 488 nm and detection at 525 nm. Dead cells were detected by excitation at 405 nm and detection with Pacific Blue Filter at 452 nm. Flow cytometry analysis was performed in BD LRS Fortessa™, by acquirement of 5.000-gated events from each sample. This figure is representative of three assays performed.

Since the percentage of FITC negative cells did not represent the real repression of transcription of GFP gene, results are shown in Mean Fluorescence Intensity (MFI) and normalized to positive control (HeLa-Tat-III/LTR/d1EGFP). It means that repression of transcription of GFP gene is reflected in a decrease of MFI.

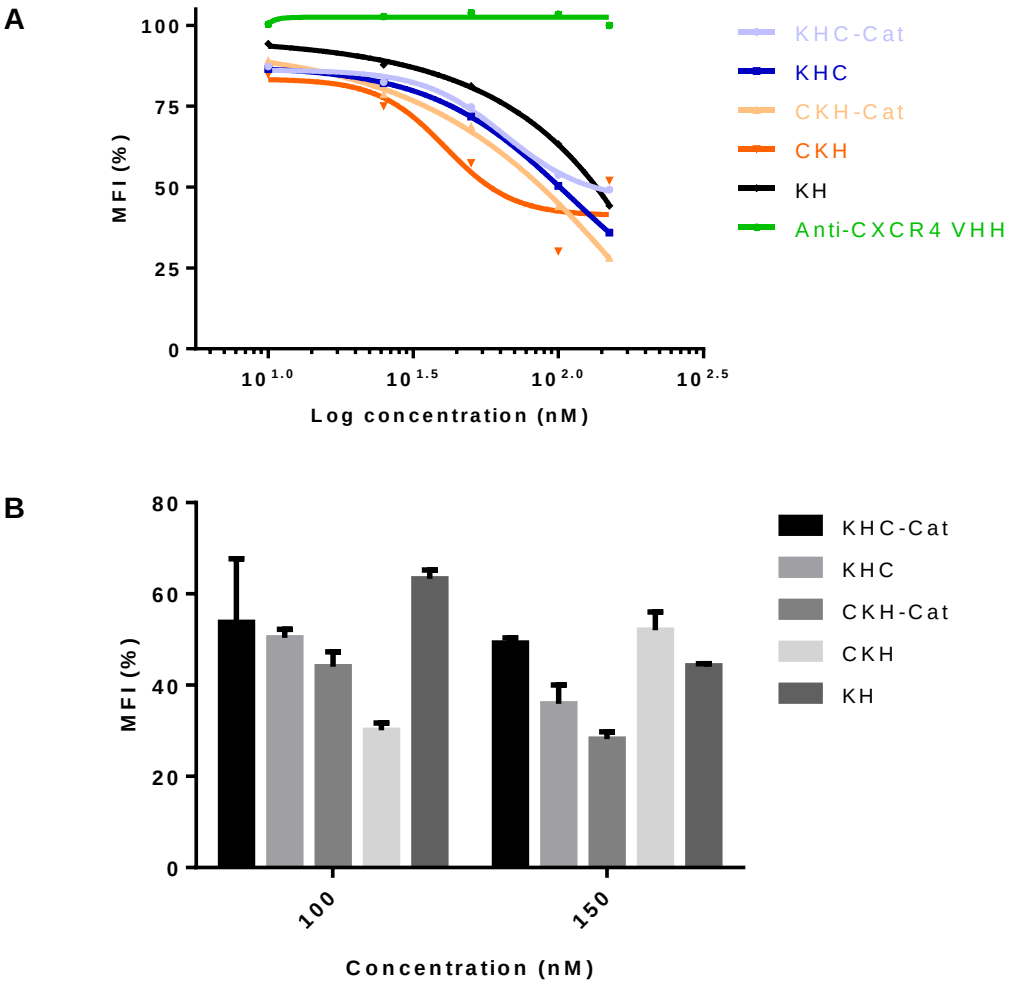


Figure 13 – Evaluation of bispecific proteins ability to repress transcription of GFP gene driven by the HIV-1 LTR promoter. Values represent mean ± SEM of the percentage of MFI obtained for three independent flow cytometry assays performed. A – A sigmoidal dose response was performed with several

concentrations of the bispecific proteins (10 nM, 25 nM, 50 nM, 100 nM and 150 nM). B – This graphical representation includes only the 100 nM and 150 nM concentrations of the bispecific proteins.

As shown in figure 13 A, CKH, CKH-Cat and KHC-Cat could repress transcription of GFP gene driven by the HIV-1 LTR, similarly to KHC and KH. Moreover, results reveal that the transcriptional repression of LTR promoter is concentration-dependent.

At 100 nM, CKH presents the highest GPF gene repression. At the same concentration, CKH-Cat presents a higher ability to repress transcription than KHC, KHC-Cat and KH (Figure 13 B). The actual number of MFI was 30,1 % for CKH, 44 % for CKH-Cat, 50,4 % for KHC, 53,8 % for KHC-Cat and 63,3 % for KH.

At 150 nM, CKH-Cat presents the highest GPF gene repression. Results for KHC and KHC-Cat were similar with KH. However, CKH which was presented the highest ability to repress transcription at 100 nM, results show that an increase in the concentration of CKH leads to a decrease of this ability and also induces cell death. The actual number of MFI was 28,2 % for CKH-Cat, 35,9 % for KHC, 44,2 % for KH, 49,2 % for KHC-Cat and 52 % for CKH.

As figure 13 A demonstrates, anti-CXCR4 VHH did not repress the reporter (~100 % of MFI), which it was expected.

3.5. Construction and Expression of Trastuzumab-E2C

For the development of an additional therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers, we constructed three bispecific proteins which contain a monoclonal antibody that interferes with the HER2 receptor (Trastuzumab, Herceptin®) fused to an artificial zinc-finger protein with the KRAB repressor domain designed to target the promoter of the protooncogene *erbB-2/HER-2* [80]. In order to facilitate the release of the zinc-finger in the target cells, in one of the constructions (Trastuzumab-E2C-Cat) cathepsin B cleavage site was introduced between the KRAB-E2C and the Trastuzumab. The same procedure was done with MMP-9 cleavage site in another construct (Trastuzumab-E2C-MMP9).

After the transfection in HEK293T cells, the supernatants were recovered and an immunoprecipitation assay was performed with a native Protein A Sepharose. To confirm protein expression in this cell line, a western blot assay was conducted with the total protein of each sample (Figure 14).

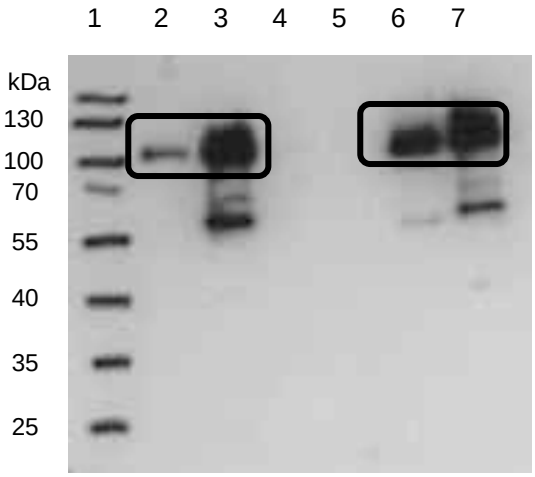


Figure 14 - Western Blot analysis of transfection of Trastuzumab-E2C, Trastuzumab-E2C-MMP-9 and Trastuzumab-E2C-CathepsinB in HEK293T cell line. Legend: 1) Protein Ladder; 2) Supernatant of Trastuzumab-E2C-Cat; 3) Lysate of Trastuzumab-E2C-Cat; 4) Supernatant of Trastuzumab-E2C-MMP9; 5) Lysate of Trastuzumab-E2C-MMP9; 6) Supernatant of Trastuzumab-E2C; 7) Lysate of Trastuzumab-E2C.

Results reveal that the fusion proteins obtained present the expected molecular weight of ~185 kDa. Although supernatants and lysates from transfection cells show detectable amounts of Transtuzumab-E2C and Trastuzumab-E2C-Cat, the proteins expression level is higher in the lysate fractions.

Regarding Trastuzumab-E2C-MMP9, this construct no present protein expression in none of the recovered fractions, which indicates that this specific position on IgG framework cannot support MMP-9 cleavage site.

DISCUSSION

Over the past two decades, gene therapy has emerged as a promising alternative to treat a variety of human diseases. The understanding of gene regulation and the structure and function of the human genome led to the emergence of new gene therapy strategies. The success of these new strategies are related with the use of an efficient, specific and non-toxic gene delivery system. One of the strategies that have been developed is the use of zinc-finger proteins. Due to their capacity to bind DNA in a specific and efficient manner, they emerged as a powerful tools for gene therapy treatment. However, several limitations must be overcome in the future, in particular their specificity to the target cells.

In order to overcome the limitations presented by zinc-finger proteins, several strategies have been assessed, in particular antibody delivery strategies. In fact, monoclonal antibodies have demonstrated enormous potential as new classes of drugs. Although conventional antibodies must overcome the critical issue of insufficient efficacy. The large molecular sizes of the conventional antibodies translates to inability to target intracellular molecules and low tissue penetration.

With this in mind, new approaches of antibody production started to appear. The emergence of new engineered antibodies for clinical applications was possible taking into account the existing knowledge of the structure and functions of antibodies. Therefore, antibody fragments emerged as a new therapeutic strategy, such as camelid VHH-single domain antibody that is the smallest antibody fragment. It is composed of only the variable domain of an antibody heavy chain, and for this reason exhibits several advantages comparing with the conventional IgGs, making them promising tools for research, diagnosis and therapy.

Taking into account the limitations of zinc-finger proteins and the therapeutic potential of antibody fragments, this thesis aims to study the improvement of a therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers previously developed in our laboratory. For this purpose, we engineered and constructed alternative bispecific proteins of an artificial zinc-finger protein against HIV-1 LTR promoter (HLTR3) [75] coupled to a CXCR4-specific nanobody (VHH) [31]. Regarding to zinc-finger protein, HLTR3 was designed to target the HIV-1 genome, in particular the HIV-1 LTR promoter. In order to repress transcription from the HIV-1 LTR promoter, HLTR3 was fused to the KRAB repressor domain [75].

To validate our strategy we designed four bispecific proteins additionally to the previously studied KHC (Cunha-Santos *et al.*, unpublished results). In one construction (KHC-Cat), the KRAB-HLTR3 was constructed on the N-terminal, similarly to KHC, while

in the three others (CKH, CKH-Cat and CKH-MMP9) was added on the C-terminal. In KHC-Cat and CKH-Cat, cathepsin B cleavage site was introduced between the KRAB-HLTR3 and the anti-CXCR4 VHH, in order to facilitate the release of the zinc-finger in the target cells [84]. For the same purpose, in CKH-MMP9, MMMP-9 cleavage site was introduced in the same position as described above. The differences between the constructions were designed to evaluate which is the best conformation and consequently which presents the better functionally.

After construction of all bispecific proteins, as explained in materials and methods section, they were subcloned into the same bacterial expression vector pET-21a (+) [91], where KHC had been previously cloned (Cunha-Santos *et al*., unpublished results), which is a high expression level. KH and KHC constructions cloned into pET-21a (+) were kindly provided by C. Cunha-Santos (Joao Gonçalves lab; unpublished results). KH was used as control for the next assays.

Afterwards, all constructions were transformed in *E.coli* BL21 (DE3) and were tested for optimal expression conditions. In this way, it was performed a small scale expression and test purification. The soluble and insoluble fractions were applied to SDS-PAGE. Unlike CKH-MMP9, the other three bispecific proteins presented high protein expression, but all in the insoluble fraction. This is an indicator that these proteins are expressed in inclusion bodies [92], similarly to what was observed previously for KH and KHC (Cunha-Santos, unpublished results), then the option is to denature and refold the proteins. Although CKH-MMP9 presented residual protein expression in both fractions, which indicates that this specific position on VHH framework cannot support MMP-9 cleavage site. This alteration affects the antibody stability/solubility. For this reason, the next assays were only performed with KHC, KHC-Cat, CKH, CKH-Cat and KH.

After the optimization of expression conditions, the next step was to purify the recombinant proteins. Since all constructions include a histidine tag (His8) on the C-terminal, they were purified by Immobilized Metal Affinity Chromatography (IMAC) with a Ni-NTA His GraviTrap™. Afterwards recombinant proteins were submitted to buffer exchange to a more suitable buffer. SDS-PAGE and Western Blot results showed a single protein band with the expected molecular weights for the proteins under reducing conditions (~55 kDa for the bispecific proteins and ~35 kDa for KH) and protein concentration was quantified by Bradford method [93]. The total amount obtained was 309 µg for KHC, 300 µg for KHC-Cat, 126 µg for CKH, 205 µg for CKH-Cat and 138 µg for

KH. Therefore, the obstacles have been overcome and the first task of this thesis was well succeeded.

After expression and purification of recombinant proteins, it was necessary to study the functionality of proteins. Preliminary test of the binding activity of the HLTR3 zinc-finger were performed by ELISA using HLTR3 binding site [75] as antigen. Results revealed that CKH, CKH-Cat and KHC-Cat bind specifically to HLTR3 binding site in a concentration-dependent manner like KHC and KH. Moreover, binding of CKH-Cat and CKH were similar with the binding of KH. Since KH was used as a positive control, this suggests that anti-CXCR4 VHH on the N-terminal does not influence the binding activity of the zinc-finger. Although these constructions presented a higher binding activity than KHC-Cat and KHC, which contain anti-CXCR4 VHH on the C-terminal. These results suggest that anti-CXCR4 on the C-terminal hampers the binding activity of the zinc-finger.

After the evaluation of proteins binding activity by ELISA, it was necessary to study the proteins specific binding to CXCR4 receptor. Since all constructions are bispecific proteins it was necessary to evaluate both function domains. CXCR4 is a transmembrane receptor and consequently it is impossible to isolate it. [87]. Since CXCR4 receptor is highly expressed in Jurkat E6-1 T-cells [94], proteins specific binding were performed in this cell line and analyzed by flow cytometry. In addition, CXCR4 receptor internalizes shortly after ligand attachment [89]. For this reason, in order to evaluate CXCR4 surface binding, the assays were performed at 4 °C.

Results revealed that KHC-Cat, CKH and CKH-Cat bind specifically to CXCR4 receptor at the surface, similarly to KHC. Unmodified anti-CXCR4 VHH, which was used as a positive control, presented the highest binding ability (41,4 % of FITC positive cells). KHC-Cat presented a higher binding ability than CKH, and also higher than KHC and CKH-Cat, probably due to conformation of the proteins (34,6 % for KHC-Cat, 21,1 % for CKH, 15,5 % for KHC and 13,7 % for CKH-Cat). In CKH and CKH-Cat, anti-CXCR4 VHH on the N-terminal may not be sufficiently exposed to achieve optimal binding to CXCR4 receptor. In resume, these results indicate that the unmodified anti-CXCR4 VHH binding ability was maintained in the recombinant proteins, in particular in KHC-Cat.

After the evaluation of both function domains of the bispecific proteins, the second task of this thesis was completed. The next step was to validate the recombinant proteins ability to repress transcription from the HIV-1 LTR promoter. For this purpose, it was assessed repression of transcription of a destabilized green fluorescent protein (GFP) reporter [90] driven by the HIV-1 LTR.

Therefore, the assay was performed in HeLa-Tat-III/LTR/d1EGFP cell line, which contains GFP gene reporter under the control of the LTR promoter. This cell line constitutively express GFP, presenting its fluorescence emission spectrum in green (520 nm) [95]. Since GFP has a half-life of about 1 h [95], cells were incubated with KHC, KHC-Cat, CKH, CKH-Cat and KH for 3 h at 37 °C. The assay was evaluated by flow cytometry, since when proteins bind to HIV-1 LTR promoter and repress transcription of GFP gene, the target cells decrease its green fluorescence emission.

To evaluate the real repression of transcription of GFP gene, results were analyzed in Mean Fluorescence Intensity (MFI) and normalized to positive control (HeLa-Tat-III/LTR/d1EGFP – 100 % of MFI). Thus, repression of transcription of GFP gene is reflected in a decrease of MFI. Results revealed that KHC-Cat, CKH and CKH-Cat could repress transcription of GFP gene driven by the HIV-1 LTR promoter in a concentration-dependent manner like KHC and KH. In fact, an increase in the protein concentration lead to a higher ability to repress transcription of GFP gene. Although several protein concentrations assessed in the assay, the GFP expression levels started to present significant alterations at 100 nM.

At 100 nM CKH-Cat presented a higher ability to repress transcription (44 % of MFI) than KHC (50,4 %) and KHC-Cat (53,8 %). As expected, with an increasing concentration, these bispecific proteins presented a higher GFP gene repression. In fact, at 150 nM CKH-Cat presented the highest GFP gene repression (28,2 %). At this concentration, KHC still presents a higher repression (35,9 %) than KHC-Cat (49,2%).

Regarding to CKH, this construct presented the highest GFP gene repression at 100 nM (30,1 %). Although at 150 nM, CKH presented a decrease in ability to repress transcription of GFP gene (52 %). In addition, this increase in concentration induced cell death by unknown mechanisms.

On the other hand, at 100 nM these three bispecific proteins showed a higher ability to repress transcription than KH (63,3 % of MFI), similarly to KHC. With an increasing concentration, CKH-Cat and KHC still show more ability to repress transcription than KH (44,2%), which is similar with KHC-Cat. This result reinforces previous results of our laboratory (Cunha-Santos, unpublished results) that when zinc-finger is coupled to an anti-CXCR4 VHH the protein delivery is more efficient, which lead to a higher ability to repress transcription of GFP gene.

In resume, CKH and CKH-Cat present more ability to repress transcription than KHC and KHC-Cat, probably due to their conformations. CKH and CKH-Cat conformations allow a higher exposure of the zinc-finger (on the C-terminal) that may

facilitate binding to the HIV-1 LTR promoter. On the other hand, in KHC and KHC-Cat the zinc-finger position on the N-terminal may not be sufficiently exposed to achieve optimal binding to the promoter, which was reflected in a smaller GFP gene repression.

Moreover, the influence of cathepsin B cleavage site in these constructions is not clear. It is expected that cathepsin B facilitates the release of the zinc-finger on the target cells and consequently the repression of transcription of GFP gene. However in this type of assays, it was not possible to evaluate the cathepsin B influence. Although the results reveal that cathepsin B cleavage site could influence the protein conformation.

In fact, CKH shows the highest ability to bind HIV-1 LTR promoter and repress transcription of GFP gene (~ 55 %) at lower concentrations (100 nM). However with an increasing concentration, this recombinant protein induces cell death. On the other hand, CKH-Cat repressed transcription of GPF gene of about ~ 45 % at lower concentrations, and no significant cytotoxicity was observed in a higher concentrations. For this reason, the choice of the bispecific protein which presents the best functionality is difficult.

According to the binding activity of the KRAB-HLTR3 zinc-finger in the different bispecific proteins (ELISA), CKH and CKH-Cat presented a higher binding activity than KHC and KHC-Cat, which is in concordance with the results of the evaluation of proteins ability to repress transcription of GFP gene. Regarding to the proteins specific binding to CXCR4 receptor, the results revealed that KHC-Cat presented the highest binding activity. These results suggest that the binding activity is important but is not essential to shutdown, which occurs downstream to binding by unknown mechanism.

The problems were overcome and the third task of this thesis was completed. We demonstrated that CKH, CKH-Cat and KHC-Cat repress the GFP expression by inhibition of the HIV-1 LTR promoter as previously proved for KHC and KH (Cunha-Santos *et al*., unpublished results). Since activation of the HIV-1 LTR promoter lead to expression of the HIV-1 genome, these results suggest that in the presence of the HIV-1 genome, these recombinant proteins inhibit the HIV-1 LTR promoter and consequently repress transcription of the HIV-1 genome. A previous study from Segal *et al.* presented similarities results with ours in the KH capacity to repress transcription from the HIV-1 LTR promoter [75].

Simultaneously, and due to the fact that antibody fragments (e.g. nanobodies) exhibit several limitations in pharmacokinetics, such as a rapid clearance, other therapeutic strategy was developed. We designed and constructed three bispecific proteins which contain a monoclonal antibody that interferes with the HER2 receptor (Trastuzumab, Herceptin®) fused to an artificial zinc-finger protein with the KRAB

repressor domain designed to target the promoter of the protooncogene *erbB-2/HER-2.* In order to facilitate the release of the zinc-finger in the target cells, in one of the constructions (Trastuzumab-E2C-Cat) cathepsin B cleavage site was introduced between the KRAB-E2C and the Trastuzumab. The same procedure was done with MMP-9 cleavage site in another construct (Trastuzumab-E2C-MMP9).

Preliminary transfection assays and western blot showed that Transtuzumab-E2C and Trastuzumab-E2C-Cat were successfully constructed and expressed in HEK293T cell line. Although Trastuzumab-E2C-MMP9 didn't present protein expression in none of the recovered fractions, which indicate that MMP-9 cleavage site plays an important role in antibody stability/solubility.

In conclusion, these bispecific proteins improve the antibody delivery strategy to gene manipulation previously developed in our laboratory. In fact, these proteins are a promising tool to be applied in the clinical.

CONCLUSIONS AND FUTURE PERSPECTIVES

Gene-based therapy is a promising alternative to treat a variety of human diseases, whether acquired or heritable. The main key to the success of the clinical gene therapy is the use of an efficient, specific and non-toxic gene delivery systems. The understanding of gene regulation and the structure and function of the human genome led to the emergence of new gene therapy strategies. Zinc-finger proteins emerged as a powerful tools for gene therapy treatment, due to their capacity to bind DNA in a specific and efficient manner. However, several limitations must be overcome in the future, in particular their specificity to the target cells.

In the presented work, we demonstrate proof-of-concept for the improvement of a previously studied therapeutic strategy of gene manipulation by antibody delivery of zinc-fingers. For this purpose, we constructed several bispecific proteins, alternatives of the previously constructed single domain antibody (VHH) that target the CXCR4 cell receptor coupled to an artificial zinc-finger protein (KRAB-HLTR3) that bind and repress transcription from HIV-1 LTR promoter.

Initially, we successfully engineered the bispecific proteins and optimized for bacterial expression and IMAC purification with high yields of purified soluble protein. Afterwards, we evaluated the proteins specificity and affinity to their targets through *in vitro* assays. We demonstrated that CKH, CKH-Cat and KHC-Cat bind specifically to HLTR3 binding site in a concentration-dependent manner, similarly to KHC and KH. We also showed that CKH, CKH-Cat and KHC-Cat bind specifically to CXCR4 receptor at the surface, similarly to KHC. Finally, we evaluate the bispecific proteins ability to gene manipulation, in particular to repress transcription from the HIV-1 LTR promoter. We assessed repression of transcription of a destabilized green fluorescent protein (GFP) reporter and we demonstrated that these alternative bispecific proteins repress transcription of GFP gene in a concentration-dependent manner like KHC and KH. Although CKH and CKH-Cat presented an even greater ability to repress transcription of GFP gene. Since activation of the HIV-1 LTR promoter lead to expression of the HIV-1 genome, these results suggest that in the presence of the HIV-1 genome, these proteins inhibit the HIV-1 LTR promoter and consequently repress transcription of the HIV-1 genome.

Regarding to Trastuzumab-E2C and Trastuzumab-E2C-Cat, preliminary assays demonstrated that these bispecific proteins were successfully constructed and expression in HEK293T.

In conclusion, results presented in this dissertation demonstrated that these therapeutic proteins improve the antibody delivery strategy to gene manipulation previously developed in our laboratory. In fact, these proteins are a promising tool to be applied in the clinical. Furthermore, these recombinant proteins can be designed and engineered to use in other therapeutic applications.

As a future perspectives, it will necessary to perform more assays in order to evaluate the bispecific proteins internalization via CXCR4. In addition, in order to validate that bispecific proteins binding and internalization are via CXCR4 receptor, the same assay must be performed in Jurkat CXCR4 negative cell line. Moreover, functional assays are necessary to evaluate the influence of cathepsin B in the release of the zinc-finger in the target cells. Regarding to bispecific proteins ability to repress transcription of the HIV-1 genome by inhibition of the LTR promoter, further *in vitro* assays have to be performed in a cell line which contains integrated HIV-1 genome. Additionally, to evaluate the repression of viral replication by these bispecific proteins, infection assays must be performed in Jurkat cell line or primary $CD4^+$ T-lymphocytes. The infection can be performed with HIV-1 laboratory-adapted strains (e.g. HIV-1_{NL4-3}) or HIV-1 primary isolates.

REFERENCES

1. Buss, N. A., Henderson, S. J., McFarlane, M., Shenton, J. M. & de Haan, L. Monoclonal Antibody Therapeutics: History And Future. *Curr. Opin. Pharmacol.* **12,** 615–22 (2012).

2. Weiner, L. M., Surana, R. & Wang, S. Monoclonal Antibodies: Versatile Platforms For Cancer Immunotherapy. *Nat. Rev. Immunol.* **10,** 317–27 (2010).

3. Shuptrine, C. W., Surana, R. & Weiner, L. M. Monoclonal antibodies for the treatment of cancer. *Semin. Cancer Biol.* **22,** 3–13 (2012).

4. Kindt, T. J., Goldsby, R. A. & Osborne, B. A. *Kuby Immunology, 6th Edition*. *New York: W. H. Freeman and Company* **6th editio,** (2007).

5. Aires, F., Corte-real, S. & Goncalves, J. Recombinant Antibodies as Therapeutic Agents: Pathways for Modeling New Biodrugs. *BioDrugs* **22,** 301–314 (2008).

6. Hansel, T. T., Kropshofer, H., Singer, T., Mitchell, J. A. & George, A. J. The Safety and Side Effects of Monoclonal Antibodies. *Nat. Rev. Drug Discov.* **9,** 325–38 (2010).

7. Brekke, O. H. & Sandlie, I. Therapeutic Antibodies For Human Diseases At The Dawn Of The Twenty-First Century. *Nat Rev Drug Discov* **2,** 52–62 (2003).

8. Kubota, T. *et al.* Engineered Therapeutic Antibodies With Improved Effector Functions. *Cancer Sci.* **100,** 1566–72 (2009).

9. Dunkelberger, J. R. & Song, W.-C. Complement And Its Role In Innate And Adaptive Immune Responses. *Cell Res.* **20,** 34–50 (2010).

10. Zipfel, P. F. & Skerka, C. Complement Regulators And Inhibitory Proteins. *Nat. Rev. Immunol.* **9,** 729–40 (2009).

11. Stoermer, K. A. & Morrison, T. E. Complement And Viral Pathogenesis. *Virology* **411,** 362–73 (2011).

12. Nimmerjahn, F. & Ravetch, J. V. Fcgamma Receptors: Old Friends And New Family Members. *Immunity* **24,** 19–28 (2006).

13. Carter, P. Improving The Efficacy Of Antibody-Based Cancer Therapies. *Nat. Rev. Cancer* **1,** 118–129 (2001).

14. Elgert, K. D. *Immunology: Understanding The Immune System, 2nd Edition*. *Willey-Blackwell* (2002).

15. Liu, J. K. H. The History Of Monoclonal Antibody Development - Progress, Remaining Challenges And Future Innovations. *Ann. Med. Surg.* **3,** 113–6 (2014).

16. Kohler, G. & Milstein, C. Continuous Cultures Of Fused Cells Secreting Antibody Of Predefined Specificity. *Nature* **256,** 495–7 (1975).

17. Jaffers, G. J. *et al.* Monoclonal Antibody Therapy. Anti-Idiotypic And Non-Anti-Idiotypic Antibodies To OKT3 Arising Despite Intense Immunosuppression. *Transplantation* **41,** 572–8 (1986).

18. Holliger, P. & Hudson, P. J. Engineered Antibody Fragments And The Rise Of Single Domains. *Nat. Biotechnol.* **23,** 1126–36 (2005).

19. Kipriyanov, S. M. Generation Of Antibody Molecules Through Antibody Engineering. *Methods Mol. Biol.* **207,** 3–25 (2003).

20. Souriau, C. & Hudson, P. J. Recombinant Antibodies For Cancer Diagnosis And Therapy. *Expert Opin. Biol. Ther.* **3,** 305–318 (2003).

21. Wörn, A. & Plückthun, A. Stability Engineering Of Antibody Single-Chain Fv Fragments. *J. Mol. Biol.* **305,** 989–1010 (2001).

22. Ahmad, Z. A. *et al.* ScFv Antibody: Principles And Clinical Application. *Clin Dev Immunol* **8,** (2012).

23. Goel, A. *et al.* Single-Chain Fv's : Novel Imaging Agents for Rapid In Vivo Localization of Human Colon Carcinoma. *J Nucl Med* **42,** 1519–27 (2001).

24. Tomlinson, I. & Holliger, P. Methods For Generating Multivalent And Bispecific Antibody Fragments. *Methods Enzymol.* **326,** 461–79 (2000).

25. Todorovska, A. *et al.* Design And Application Of Diabodies, Triabodies And Tetrabodies For Cancer Targeting. *J. Immunol. Methods* **248,** 47–66 (2001).

26. Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T. & Winter, G. Binding Activities Of A Repertoire Of Single Immunoglobulin Variable Domains Secreted From Escherichia Coli. *Nature* **341,** 544–546 (1989).

27. Hamers-Casterman, C Atarhouch, T. *et al.* Naturally Occurring Antibodies Devoid Of Light Chains. *Nature* **363,** 446–8 (1993).

28. Dooley, H. & Flajnik, M. F. Shark Immunity Bites Back: Affinity Maturation And Memory Response In The Nurse Shark, Ginglymostoma Cirratum. *Eur. J. Immunol.* **35,** 936–45 (2005).

29. Kolkman, J. A. & Law, D. A. Nanobodies - From Llamas To Therapeutic Proteins. *Drug Discov. Today Technol.* **7,** e95–e146 (2010).

30. Wesolowski, J. *et al.* Single Domain Antibodies: Promising Experimental And Therapeutic Tools In Infection And Immunity. *Med Microbiol Immunol* **198,** 157–174 (2009).

31. Jähnichen, S. *et al.* CXCR4 Nanobodies (VHH-Based Single Variable Domains) Potently Inhibit Chemotaxis And HIV-1 Replication And Mobilize Stem Cells. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 20565–70 (2010).

32. De Marco, A. Biotechnological Applications Of Recombinant Single-Domain Antibody Fragments. *Microb. Cell Fact.* **10,** 44 (2011).

33. Harmsen, M. M. & De Haard, H. J. Properties, Production, And Applications Of Camelid Single-Domain Antibody Fragments. *Appl Microbiol Biotechnol* **77,** 13–22 (2007).

34. Desmyter, A., Decanniere, K., Muyldermans, S. & Wyns, L. Antigen Specificity And High Affinity Binding Provided by One Single Loop of a Antigen Specificity and High Affinity Binding Provided by One Single Loop of a Camel Single-Domain Antibody. *J. Biol. Chem.* **276,** 26285–90 (2001).

35. Saerens, D., Ghassabeh, G. H. & Muyldermans, S. Single-Domain Antibodies As Building Blocks For Novel Therapeutics. *Curr. Opin. Pharmacol.* **8,** 600–8 (2008).

36. Birch, J. R. & Racher, A. J. Antibody Production. *Adv. Drug Deliv. Rev.* **58,** 671–685 (2006).

37. Kozlowski, S. & Swann, P. Current And Future Issues In The Manufacturing And Development Of Monoclonal Antibodies. *Adv. Drug Deliv. Rev.* **58,** 707–22 (2006).

38. Emery, D. W. Gene Therapy For Genetic Diseases: On The Horizon. *Clin. Appl. Immunol. Rev.* **4,** 411–422 (2004).

39. Smith, K. R. Gene Therapy: Theoretical and Bioethical Concepts. *Arch Med Res* **34,** 247–268 (2003).

40. Mhashilkar, A., Chada, S., Roth, J. A. & Ramesh, R. Gene Therapy. Therapeutic Approaches And Implications. *Biotechnol Adv* **19,** 279–297 (2001).

41. Ginn, S. L., Alexander, I. E., Edelstein, M. L., Abedi, M. R. & Wixon, J. Gene Therapy Clinical Trials Worldwide to 2012 - an update. *J. Gene Med.* **15,** 65–77 (2013).

42. Young, L. S., Searle, P. F., Onion, D. & Mautner, V. Viral Gene Therapy Strategies: From Basic Science To Clinical Application. *J Pathol* **208,** 299–318 (2006).

43. Mountain, A. Gene Therapy: The First Decade. *Trends Biotechonol* **18,** 119–128 (2000).

44. Gersbach, C. A., Gaj, T. & Barbas III, C. F. Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. *Acc Chem Res* **47,** 2309–2318 (2014).

45. Gaj, T., Guo, J., Kato, Y., Sirk, S. J. & Barbas III, C. F. Targeted Gene Knockout By Direct Delivery of ZFN Proteins. *Nat. Methods* **9,** 805–807 (2013).

46. Liu, J., Gaj, T., Wallen, M. C. & Barbas III, C. F. Improved Cell-Penetrating Zinc-Finger Nuclease Proteins for Precision Genome Engineering. *Mol Ther Nucleic Acids* **4,** (2015).

47. Gaj, T., Liu, J., Anderson, K. E., Sirk, S. J. & Barbas III, C. F. Protein Delivery Using Cys2-His2 Zinc-Finger Domains. *ACS Chem Biol* **9,** 1662–1667 (2014).

48. Nelson, D. L. & Cox, M. M. *Lehinnger. Principles of Biochemistry. 5th Edition*. (2008).

49. Klug, A. Towards Therapeutic Applications Of Engineered Zinc Finger Proteins. *Febs Lett* **579,** 892–894 (2005).

50. Siegal, F. P. *et al.* Severe Acquired Immunodeficiency In Male Homosexuals, Manifested By Chronic Perianal Ulcerative Herpes Simplex Lesions. *N. Engl. J. Med.* **305,** 1439–1444 (1981).

51. Gottlib, M. S. *et al.* Pneumocystis Carinii Pneumonia And Mucosal Candidiasis In Previously Healthy Homosexual Man: Evidence Of a New Acquired Cellular Immunodeficiency. *N. Engl. J. Med.* **305,** 1425–1431 (1981).

52. Masur, H. *et al.* An Outbreak Of Commnity-Acquired Pneumocystis Carinii Pneumonia: Initial Manifestation Of Cellular Immune Dysfunction. *N. Engl. J. Med.* **305,** 1431–8 (1981).

53. Gilbert, P. B. *et al.* Comparison Of HIV-1 And HIV-2 Infectivity From A Prospective Cohort Study In Senegal. *Stat. Med.* **22,** 573–93 (2003).

54. Marlink, R. *et al.* Reduced Rate of Disease Development After HIV-2 Infection as Compared to HIV-1. *Science (80-.).* **265,** 1587–90 (1994).

55. Gallo, R. *et al.* Isolation of Human T-cell Leukemia Virus In Acquired Immune Efficiency Syndrome (AIDS). *Science (80-.).* **220,** 865–7 (1983).

56. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic Retrovirus From a Patient at Risk For Acquired Immune Deficiency Syndrome (AIDS). *Science (80-.).* **220,** 868–71 (1983).

57. Ho, D. D. *et al.* Rapid Turnover Of Plasma Virions And CD4 Lymphocytes In HIV-1 Infection. *Nature* **373,** 123–126 (1995).

58. Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science (80-.).* **271,** 1582–1586 (1996).

59. Weiss, R. A. HIV Receptors and Cellular Tropism. *IUBMB Life* **53,** 201–205 (2002).

60. Richman, D. D. *et al.* The Challenge of Finding a Cure for HIV Infection. *Science (80-.).* **323,** 1304–1307 (2009).

61. Turner, B. G. & Summers, M. F. Structural Biology Of HIV. *J. Mol. Biol.* **285,** 1–32 (1999).

62. Wagner, E. K., Hewlett, M. J., Bloom, D. C. & Camerini, D. *In Basic Virology. 3th Edition.* (2007).

63. HIV Sequence Database. at <http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html>

64. Costin, J. M. Cytopathic Mechanisms of HIV-1. *Virol J* **4,** 100 (2007).

65. Frankel, A. D. & Young, J. A. T. HIV-1 : Fifteen Proteins And An RNA. *Annu Rev Biochem* **67,** 1–25 (1998).

66. Robinson, H. L. New Hope For An AIDS Vaccine. *Nat. Rev. Immunol.* **2,** 239–50 (2002).

67. Sherman, M. P. & Greene, W. C. Slipping Through The Door: HIV Entry Into The Nucleus. *Microbes Infect* **4,** 67–73 (2002).

68. Gorry, P. R. & Ancuta, P. Coreceptors And HIV-1 Pathogenesis. *Curr HIV/AIDS Rep* **8,** 45–53 (2011).

69. Clapham, P. R. & Mcknight, Á. HIV-1 Receptors And Cell Tropism. *Br. Med. Bull.* **58,** 43–59 (2001).

70. Shah, V. B. & Aiken, C. HIV Nuclear Entry: Clearing the Fog. *Viruses* **2,** 1190–1194 (2010).

71. Bukrinsky, M. A Hard Way to the Nucleus. *Mol Med* **10,** 1–5 (2004).

72. Laskey, S. B. & Siliciano, R. F. A Mechanistic Theory To Explain The Efficacy Of Antiretroviral Therapy. *Nat. Rev. Microbiol.* **12,** 772–780 (2014).

73. Strayer, D. S. *et al.* Current Status of Gene Therapy Strategies to Treat HIV/AIDS. *Mol Ther* **11,** 823–842 (2005).

74. Papworth, M., Kolasinska, P. & Minczuk, M. Designer Zinc-Finger Proteins And Their Applications. *Gene* **366,** 27–38 (2006).

75. Segal, D. J. *et al.* Attenuation of HIV-1 Replication In Primary Human Cells With a Designed Zinc Finger Transcription Factor. *J. Biol. Chem.* **279,** 14509–19 (2004).

76. Reynolds, L. *et al.* Repression Of The HIV-1 5'LTR Promoter And Inhibition Of HIV-1 Replication By Using Engineered Zinc-Finger Transcription Factors. *Proc Natl Acad Sci USA* **100,** 1615–20 (2003).

77. Kim, Y. *et al.* Artificial Zinc Finger Fusions Targeting Sp1-binding Sites and The Trans - Activator-Responsive Element Potently Repress Transcription and Replication of HIV-1. *J Biol Chem* **280,** 21545–21552 (2005).

78. Eberhardy, S. R. *et al.* Inhibition of Human Immunodeficiency Virus Type 1 Replication with Artificial Transcription Factors Targeting the Highly Conserved Primer-Binding Site. *J. Virol.* **80,** 2873–2883 (2006).

79. Cunha-santos, C., Figueira, T. N., Borrego, P., Oliveira, S. S. & Gonçalves, J. Development of Synthetic Light-Chain Antibodies as Novel and Potent HIV Fusion Inhibitors. *AIDS* **30,** 1691–701 (2016).

80. Beerli, R., Segal, D., Dreier, B. & Barbas III, C. Toward Controlling Gene Expression At Will: Specific Regulation of The ErbB-2/HER-2 Promoter By Using Polydactyl Zinc Finger Proteins Constructed From Modular Building Blocks. *Proc Natl Acad Sci USA* **95,** 14628–14633 (1998).

81. Li, X. *et al.* Site-Specific Dual Antibody Conjugation via Engineered Cysteine and Selenocysteine Residues. *Bioconjug. Chem.* **26,** 2243–8 (2015).

82. Patterson, J. T., Asano, S., Li, X., Rader, C. & Barbas III, C. F. Improving the Serum Stability of Site-Specific Antibody Conjugates with Sulfone Linkers. *Bioconjug. Chem.* **25,** 1402–7 (2014).

83. Jordan, M., Schallhorn, A. & Wurm, F. M. Transfecting Mammalian Cells: Optimization of Critical Parameters Affecting Calcium-Phosphate Precipitate Formation. *Nucleic Acids Res* **24,** 596–601 (1996).

84. Mort, J. S. & Buttle, D. J. Cathepsin B. *Int J Biochem Cell Biol* **29,** 715–720 (1997).

85. Turk, V. *et al.* Cysteine Cathepsins: From Structure, Function and Regulation To New Frontiers. *Biochim. Biophys. Acta* **1824,** 68–88 (2012).

86. Wiranowska, M., Rojiani, A. M. & Rojiani, M. V. Matrix Metalloproteinases-Modulating the Tumor Microenvironment. *J Carcinog Mutagen* **6,** 1–7 (2015).

87. Debnath, B., Xu, S., Grande, F., Garofalo, A. & Neamati, N. Small Molecule Inhibitors of CXCR4. *Theranostics* **3,** 47–75 (2013).

88. Sodhi, A., Montaner, S. & Gutkind, J. S. Viral Hijacking of G-Protein-Coupled-Receptor Signalling Networks. *Nat Rev Mol Cell Biol* **5,** 998–1012 (2004).

89. Busillo, J. M. & Benovic, J. L. Regulation Of CXCR4 Signaling. *Biochim. Biophys. Acta* **1768,** 952–963 (2007).

90. Li, X. *et al.* Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *J Biol Chem* **273,** 34970–34975 (1998).

91. Addgene. at <https://www.addgene.org/vector-database/2549/>

92. European Molecular Biology Laboratory. at <https://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/improving_protein_solubility/>

93. Bradford, M. M. A Rapid and Sensitive Method For The Quantitation of Microgram Quantities of Protein Utilizing The Principle of Protein-Dye Binding. *Anal. Biochem.* **72,** 248–254 (1976).

94. Majka, M. *et al.* Biological Significance of The Expression of HIV-Related Chemokine Coreceptors (CCR5 and CXCR4) and Their Ligands By Human Hematopoietic Cell Lines. *Leukemia* **14,** 1821–1832 (2000).

95. Chalfie, M. & Kain, S. R. *Green Fluorescent Protein: Properties, Applications and Protocols. 2nd Edition.* (2005).

ANNEXES

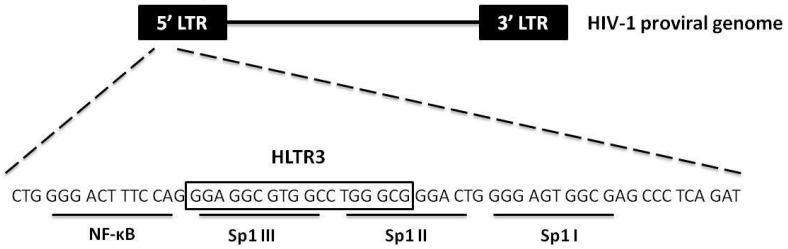


Figure 15 – Binding site of zinc-finger protein HLTR3 on HIV-1 genome. [75]

Table 2 – Primers sequence used in PCR reactions.

Primer name	Sequence 5'-3'
Primer 1	CTAGCTAGCGAGGTGCAGCTGGTGGAG
Primer 2	GCTGAGCTCTCCGCTCGTCAACGTTCTTGGTATTTTGCTTCCG CTCACGGTCACCTGGGTGC
Primer 3	GACTAGTCCCCTGAAGCCCGCCAAGAGCGCCAG AAGCAGCGGAGGAGGCGGAAGC
Primer 4	CCGCTCGAGTCATGCGTAATCAGGCACGTCGTAGGGGTACGATCCA TGGTGATGGTGATGGTGATGGTGGCTGCCTCCGCCTCCACTGCTGC TCACGGTCACCTG
Primer 5	AGAAATCACTAAGCTTCGGAATCGATGCCAAGAGCCTGACC
Primer 6	ACCCGGAGACAAGCTTCAAGAAGCGTAGTCCGGAACG
Primer 7	AGAAATCACTAAGCTCTCGACCACTGAAGCCCGC
Primer 8	AGAAATCACTAAGCAAAATACCAAGAACGTTGACGAGCG GCGGAATCTTCGGAATCGATGCCAAGAGCCTGACC

Table 3 – PCR conditions program used in PCR reactions.

Phusion Green High-Fidelity DNA Polymerase			
Cycle	**Step**	**Temperature**	**Time**
1	Initial denaturation	98 °C	30 sec.
30	Denaturation	98°C	10 sec.
	Annealing	60°C	30 sec.
	Extension	72°C	30 sec.
1	Final extension	72°C	10 min.

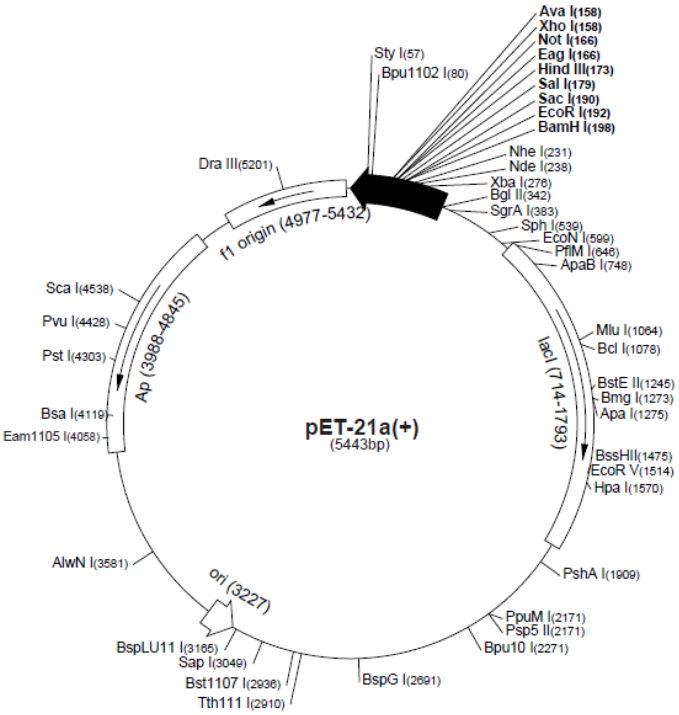
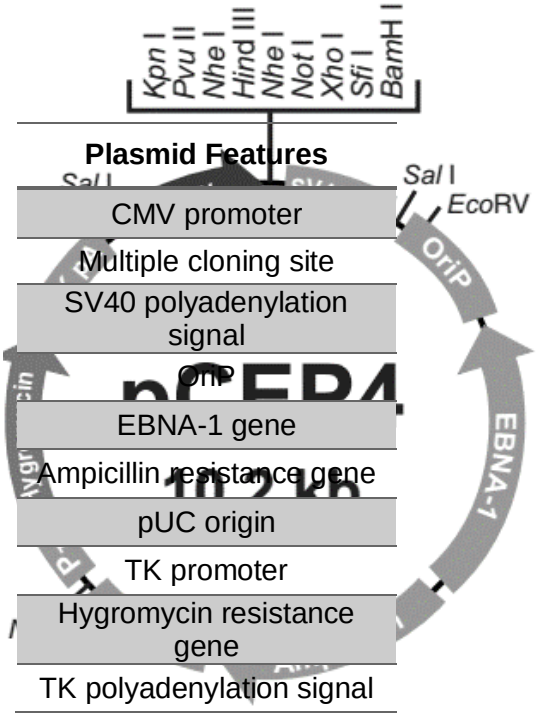


Figure 16 – Plasmid features and genomic map of pET-21a (+).

Plasmid Features
T7 promoter
T7 transcription start
T7-tag coding sequence
Multiple cloning sites
His-tag coding sequence
T7 terminator
lacI coding sequence
pBR322 origin
bla coding sequence
f1 origin



Plasmid Features
CMV promoter
Multiple cloning site
SV40 polyadenylation signal
OriP
EBNA-1 gene
Ampicillin resistance gene
pUC origin
TK promoter
Hygromycin resistance gene
TK polyadenylation signal

Figure 17 - Plasmid features and genomic map of pCEP4.